



Introdução à **técnica** de **CRISPR**

Tiago Campos Pereira
(organizador)

SBG Sociedade
Brasileira de
Genética

Ribeirão Preto, 2016

© 2016

Todos os direitos desta edição são reservados à Sociedade Brasileira de Genética.

Comissão Editorial Sociedade Brasileira de Genética

Editor

Élgion Lúcio Silva Loreto

Universidade Federal de Santa Maria

Comissão Editorial

Carlos Frederico Martins Menck

Universidade de São Paulo

Louis Bernard Klaczko

Universidade Estadual de Campinas

Marcio de Castro Silva-Filho

Universidade de São Paulo

Maria Cátira Bortolini

Universidade Federal do Rio Grande do Sul

Marcelo dos Santos Guerra Filho

Universidade Federal de Pernambuco

Pedro Manoel Galetti Junior

Universidade Federal de São Carlos

Introdução à técnica de CRISPR/Tiago Campos Pereira (organizador).
– Ribeirão Preto : Sociedade Brasileira de Genética, 2016.

250 p.

ISBN 978-85-89265-24-9

1. CRISPR. 2. Edição genômica. 3. Engenharia genética.
4. Nocauteamento gênico. I. Pereira, Tiago Campos, org.

Capa, projeto gráfico e diagramação

editora  cubo
soluções para o universo acadêmico



Sociedade
Brasileira de
Genética

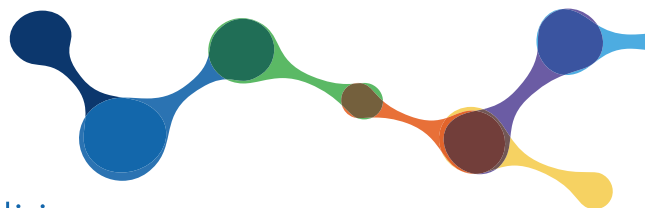
Rua Cap. Adelmio Norberto da Silva, 736

14025-670 - Ribeirão Preto - SP

16 3621-8540 | 16 3621-3552

Esta obra conta com a parceria de:

MERCK



CRISPR-Cas9 Genome Editing

Increase efficiency of genome editing using the Alt-R™ CRISPR-Cas9 System.

Now powered by the new S.p. Cas9 Nuclease 3NLS!

The Alt-R CRISPR-Cas9 System includes all of the reagents needed for successful genome editing. Based on the natural *S. pyogenes* CRISPR-Cas9 system, the Alt-R CRISPR-Cas9 System offers numerous advantages over alternative methods:

- Higher on-target potency than other CRISPR systems;
- Precise control of editing complex delivery with Cas9 ribonucleoprotein;
- Efficient delivery of the RNP with lipofection or electroporation;
- No toxicity or innate immune response activation as observed with *in vitro* transcribed Cas9 mRNA and sgRNAs.

New! Alt-R CRISPR crRNAs with increased nuclease resistance

IDT has updated the Alt-R CRISPR-Cas9 System with the addition of proprietary chemical modifications to the Alt-R CRISPR crRNA. These modifications protect the crRNA from degradation by cellular RNases, and further improve on-target editing performance. The modifications are included automatically to the final Alt-R CRISPR crRNA oligonucleotide sequence and do not require any changes to the ordering process.

For more information, ordering tool, and tutorials/video, visit www.idtdna.com/CRISPR

For Research Use Only. Not for use in diagnostic procedures.

CRISPR triggers

Product	Size	Catalog #
Alt-R™ CRISPR crRNA	2, 10 nmol tubes; 2 nmol plates	Order at www.idtdna.com/CRISPR
	5 nmol	1072532
Alt-R™ CRISPR tracrRNA	20 nmol	1072533
	100 nmol	1072534

Cas9 endonuclease

Product	Catalog #
Alt-R™ S.p. Cas9 Nuclease 3NLS (100 µg)	1074181
Alt-R™ S.p. Cas9 Nuclease 3NLS (500 µg)	1074182

Control kits*

Product	Catalog #
Alt-R™ CRISPR Control Kit, Human (2 nmol)	1072554
Alt-R™ CRISPR Control Kit, Mouse (2 nmol)	1072555
Alt-R™ CRISPR Control Kit, Rat (2 nmol)	1072556

Control kit contents

- Alt-R™ CRISPR HPRT Positive Control crRNA
- Alt-R™ CRISPR Negative Control crRNA #1
- Alt-R™ CRISPR tracrRNA
- Alt-R™ HPRT PCR Primer Mix
- Nuclease-Free Duplexing buffer

* Control kit components are also available individually.

Os autores

Andréa Almeida Carneiro é pesquisadora da Embrapa Milho e Sorgo, atuando na área de Cultura de Tecidos e Transformação Genética de Milho e Sorgo. Bióloga pela Pontifícia Universidade Católica de Minas Gerais. Possui Mestrado em Microbiologia Agrícola pela Universidade Federal de Viçosa e Doutorado em *Plant Science* pela University of Arizona, EUA.

Andréia Prata Vieira é bióloga bacharel e licenciada pela Universidade de São Paulo. Faz Doutorado Direto pelo Programa de Dupla Diplomação Internacional de Biologia Celular e Molecular de Plantas pela Universidade de São Paulo e Rutgers University – The State University of New Jersey, nos Estados Unidos.

Arthur Casulli de Oliveira é mestrando pelo Programa de Pós-Graduação de Genética da UNESP – Botucatu, no Laboratório de Genômica e Evolução Molecular (LGEM). Bacharel e licenciado em Ciências Biológicas pela UNESP – Botucatu, com estágio realizado em Earlham Institute (Norwich, Inglaterra).

Augusto Schrank é professor titular do Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, e pesquisador do Centro de Biotecnologia da UFRGS. Doutorado pela University of Manchester Institute of Science and Technology – UMIST (Manchester, Inglaterra).

Bárbara Costa Peixoto é bióloga pelo Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro. Bolsista FAPERJ Nota 10 pelo Programa de Mestrado do Instituto Nacional de Câncer.

Bruno Henrique Garcia é graduando no curso de Ciências Biológicas com ênfase em Biologia Molecular e Tecnológica pela Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP) da Universidade de São Paulo (USP).

Bruno Trevenzoli Favero é pesquisador visitante na Faculdade de Ciências da Universidade de Copenhague. Engenheiro agrônomo pela USP/ESALQ, com Mestrado em Agricultura Tropical e Subtropical pelo Instituto Agronômico de Campinas e Doutorado em Ciências Biológicas (Botânica) pela UNESP, com período sanduíche na Universidade Estadual da Carolina do Norte (NCU), Estados Unidos. Possui Pós-doutorado em Biotecnologia pela Universidade de Copenhague, Dinamarca.

Camila T. França é biomédica pela Universidade Federal de Pernambuco (UFPE), possui Mestrado em Ciências Biológicas pela UFPE e Doutorado em Imunoepidemiologia pela Universidade de Melbourne, Austrália. Atualmente pós-doutoranda entre o Walter & Eliza Hall Institute em Melbourne, Austrália, e o Pasteur Institut em Paris, França.

Carlos Eduardo Ambrósio é docente de Morfologia no Departamento de Medicina Veterinária, da Faculdade de Zootecnia e Engenharia de Alimentos da USP de Pirassununga. Médico veterinário pela UNIFEOB, com Mestrado e Doutorado em Anatomia pela FMVZ/USP. Possui Pós-doutorado em Biotecnologia de Células-Tronco pela Universidade de São Paulo.

Caroline Junqueira é pesquisadora do Centro de Pesquisas René Rachou/Fiocruz em Belo Horizonte. Formada em Ciências Biológicas pela PUC-MG, com Mestrado e Doutorado em Bioquímica e Imunologia pela UFMG. Possui Pós-doutorado em Imunologia pelo CPqRR/Fiocruz e em Biologia Celular pela Universidade de Notre Dame, EUA.

Cesar Augusto Speck Hernandez é biólogo pela Universidad del Tolima (Colômbia) e atualmente mestrando no Programa de Pós-graduação em Imunologia Básica e Aplicada da Universidade de São Paulo em Ribeirão Preto.

Cristina Bonorino é professora titular de Imunologia da PUC do Rio Grande do Sul (PUC-RS) em Porto Alegre. Bióloga pela UFRGS, com Mestrado e Doutorado em Genética e Biologia Molecular pela mesma instituição e pela Universidade do Colorado, EUA. Possui Pós-doutorado em Imunologia Celular e Molecular nas Universidades de Minnesota (UofM) e Califórnia, em San Diego (UCSD).

Daiana Renck realiza Pós-doutorado na Universidade Católica do Rio Grande do Sul (PUC-RS). Farmacêutica com ênfase em Farmácia Industrial pela PUC-RS, Mestrado e Doutorado em Biologia Celular e Molecular.

Danillo Pinhal é docente de Genética e Evolução e chefe do Departamento de Genética da UNESP – Botucatu. Biólogo pela UNESP, com Mestrado em Biologia Molecular e Estrutural, Doutorado e Pós-doutorado em Genética pela mesma instituição, com período sanduíche na Nova Southeastern University e na Vanderbilt, EUA.

Erika Kague é pesquisadora Jovem Talento no Departamento de Genética e Biologia Evolutiva da Universidade de São Paulo (USP) em São Paulo. Bióloga pela UNESP, com Doutorado em Genética e Biologia Evolutiva pela USP, com período sanduíche na Johns Hopkins University, Estados Unidos. Possui Pós-doutorado em Biologia de Desenvolvimento pela Universidade da Pennsylvania (UPENN), Estados Unidos.

Gabriel Viliod Vieira, graduado em Ciências Biológicas pelo Centro Universitário Barão de Mauá, com Mestrado em andamento em Biologia Celular e Molecular e Bioagentes Patogênicos pela Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (USP).

Gabriela Molinari Roberto é bióloga formada na Universidade de São Paulo em Ribeirão Preto e atualmente mestrando pelo Programa de Pós-graduação em Oncologia Clínica, Terapia Celular e Células-Tronco da mesma instituição. Durante a graduação complementou seus estudos nos Estados Unidos, na instituição Roanoke College e na Universidade de Stanford.

Geraldo Aleixo Passos é docente de Genética e Biologia Molecular na USP em Ribeirão Preto desde 1988. Biólogo pela UNESP (1983), Doutorado em Bioquímica (1988) pela USP e Pós-doutorado pelo Institut de Génétique Moléculaire de Montpellier, França (1994). Trabalha com expressão gênica transcricional de células do sistema imune humano e murino.

Giulliana Augusta Rangel Gonçalves é assistente de pesquisa sênior do Instituto Israelita de Pesquisa Albert Einstein (IIEP) em São Paulo. Biomédica pela UMC, com Mestrado em Biotecnologia pela UNIFESP. Possui MBA em Gestão Empresarial pela Universidade Presbiteriana Mackenzie, São Paulo.

Gleidson Silva Teixeira é pesquisador na GranBio Celere e pesquisador colaborador na Universidade Estadual de Campinas (UNICAMP), associado ao Departamento de Genética, Evolução e Bioagentes. É biólogo, Mestrado e Doutorado em Genética e Biologia Molecular pela UNICAMP.

Gonçalo Pereira é professor titular do Instituto de Biologia da Universidade Estadual de Campinas (UNICAMP) e cientista-chefe da GranBio.

Jayme A. Souza-Neto é biólogo (UENF), Doutorado em Genética (UNESP), com Pós-doutorado em Microbiologia Molecular e Imunologia pela Johns Hopkins Bloomberg School of Public Health. É docente da UNESP de Botucatu, lotado na Faculdade de Ciências Agrônômicas e no Instituto de Biotecnologia.

Joan Barau realiza Pós-doutorado como integrante do grupo de Epigenética e Biologia do Desenvolvimento no Instituto Curie em Paris, França. Biólogo pela UNICAMP, com Doutorado em Genética e Biologia Molecular pela mesma instituição.

Juliana Borsoi é doutoranda em Genética no Instituto de Biociências da Universidade de São Paulo (USP). Possui graduação em Ciências Biológicas pela mesma instituição, com período sanduíche na University of Glasgow (Escócia, Reino Unido).

Julio Cesar Cetrulo Lorenzi é pós-doutorando no laboratório de Imunologia Molecular na Rockefeller University (EUA). Biólogo formado pela Unifal-MG, com Mestrado em Imunologia Básica e Aplicada e Doutorado em Genética, ambos pela FMRP-USP.

Kaio Cesar Simiano Tavares é professor e pesquisador da UNIFOR. Possui graduação em Medicina Veterinária pela UDESC e Mestrado em Ciência Animal pela mesma instituição. Doutorado em Biotecnologia pela Rede Nordeste de Biotecnologia (RENORBIO), com período sanduíche no Instituto Roslin, da Universidade de Edimburgo, Reino Unido.

Katiuchia Uzzun Sales, formada em Odontologia, obteve o Doutorado em Patologia Bucal pela Faculdade de Odontologia da USP e especialização em Endodontia pela Faculdade de Odontologia da Universidade Santa Cecília. Realizou Pós-doutorado no National Institute of Dental and Craniofacial Research/National Institutes of Health e atualmente professora doutora pela Faculdade de Medicina de Ribeirão Preto da USP.

Larissa Cotrim de Sousa é graduada em Ciências Biomédicas pelo Instituto de Biociências da UNESP – Botucatu. Atualmente é aluna de Mestrado no Laboratório de Imunogenética Molecular da Faculdade de Medicina de Ribeirão Preto da USP.

Leonardo Chicaybam é tecnologista na área de Imunologia da Vice-Presidência de Pesquisa e Laboratórios de Referência (VPPLR) – Fiocruz/RJ. Possui Mestrado em Oncologia pelo Instituto Nacional de Câncer (INCA).

Letícia Magalhães Arruda atualmente integra a equipe do Laboratório de Biologia Sistêmica e Sintética na Faculdade de Medicina da USP de Ribeirão Preto. É bacharel em Bioquímica pela Universidade Federal de Viçosa (UFV), Doutorado em Bioquímica pela Faculdade de Medicina da USP de Ribeirão Preto. Possui Pós-doutorado em Biologia Sintética pela FMRP-USP.

Luciana Rodrigues Carvalho Barros integra o Programa de Desenvolvimento Institucional do Instituto Nacional de Câncer (INCA). Bióloga pela UFF, com Mestrado e Doutorado em Biologia Celular e Molecular pelo Instituto Oswaldo Cruz (IOC-Fiocruz) no Rio de Janeiro, com período sanduíche na Université Pierre et Marie Curie em Paris, França. Possui Pós-doutorado em Oncologia pelo INCA.

Lygia da Veiga Pereira é professora titular e chefe do Laboratório Nacional de Células-Tronco Embrionárias da Universidade de São Paulo (USP). Física pela PUC-RJ, possui Mestrado em Biofísica pela UFRJ e PhD em Genética Humana pela City University of New York/Mount Sinai Graduate School.

Marcelo C. Bassalo é doutorando na University of Colorado Boulder, Estados Unidos. Biólogo pela UNICAMP, com Mestrado em Genética e Biologia Molecular pela mesma instituição.

Maria Betânia Melo de Oliveira é bióloga pela UFRPE, Mestrado em Biologia Vegetal e Doutorado em Ciências Biológicas pela UFPE. Atualmente é professora associada I da UFPE, pesquisadora colaboradora do Centro de Pesquisas Aggeu Magalhães e líder do grupo de pesquisa NuBIOMA (Núcleo de Biossegurança e Meio Ambiente) da UFPE.

Maria Paloma Silva de Barros é bióloga pela UFPE, Mestrado em Saúde Pública pelo Centro de Pesquisas Aggeu Magalhães – Fiocruz/PE e Doutorado em Genética pela UFPE. Desenvolve parcerias com o Departamento de Microbiologia do CPqAM – Fiocruz/PE.

María Sol Brassesco Annichini é docente de Genética na Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo (FFCLRP-USP) em Ribeirão Preto. Geneticista pela UNaM (Argentina). Possui Mestrado e Doutorado em Genética pela Faculdade de Medicina de Ribeirão Preto da USP e Pós-doutorado em Oncologia Pediátrica pela mesma instituição.

Marie-Anne Van Sluys é professora titular do Departamento de Botânica da Universidade de São Paulo. Graduou-se em Ciências Biológicas pela Universidade do Estado do Rio de Janeiro (UERJ), com Doutorado pela Université Paris-Sud 11.

Martin Bonamino é pesquisador do Instituto Nacional de Câncer e especialista da Fundação Oswaldo Cruz (RJ). Biomédico pela UFRJ, com Doutorado em Química Biológica pela mesma instituição.

Natalia Nardelli Gonçalves é pós-doutoranda na área de Edição Gênica e Células-Tronco na Universidade de São Paulo (FZEA). Bióloga pela Unifeob, com Mestrado em Ciências pela Universidade de São Paulo (CENA/ESALQ), Doutorado em Biologia Celular e Molecular pelo Programa de Anatomia da FMVZ (USP), com período sanduíche na Universidade da Carolina do Norte (NCSU).

Nerry Tatiana Cecílio é Pós-doutoranda em Farmacologia pela Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo (USP). Bióloga pelo Centro Universitário Barão de Mauá, com Mestrado e Doutorado em Biologia Celular e Molecular pela Faculdade de Medicina de Ribeirão Preto da USP.

Newton Portilho Carneiro é biólogo pela Pontifícia Universidade Católica de Minas Gerais, com Mestrado em Genética e Melhoramento pela UFV e Doutorado em Plant Sciences pela University of Arizona. Atualmente é pesquisador da Empresa Brasileira de Pesquisa Agropecuária, Embrapa Milho e Sorgo (MG).

Nilza Maria Diniz é docente de Genética e de Bioética da Universidade Estadual de Londrina (UEL). Bióloga pela FFCLRP-USP, com Mestrado pela FMRP-USP e Doutorado em Genética pela mesma instituição/Washington State University, EUA. Possui Pós-doutorado em Bioética pela Cátedra UNESCO em Bioética da UnB.

Oliver Kemp é bacharel em Biotecnologia pela Universidade de Copenhagen. Atualmente é aluno de Mestrado do Programa de Pós-graduação em Biotecnologia pela mesma instituição.

Patrícia Renck Nunes faz Doutorado em Molecular Life Sciences na Escola Politécnica Federal de Lausanne, na Suíça, com bolsa da fundação Boehringer Ingelheim Fonds, na Alemanha. Biomédica pela Universidade Federal do Rio Grande do Sul, com Mestrado em Biologia Médica pela Universidade de Lausanne.

Rafael Sanguinetti Czepielewski é pós-doutorando do Laboratório de Imunologia Celular e Molecular e de Drug Discovery do Instituto de Pesquisas Biomédicas (IPB) da PUC-RS em Porto Alegre. Biólogo pela PUC-RS, com Mestrado e Doutorado em Biologia Celular e Molecular pela mesma instituição, com período sanduíche na Icahn School of Medicine at Mount Sinai, Nova York, EUA.

Rafaela de Freitas Martins Felício é biomédica pela Universidade Federal do Mato Grosso (UFMT) no campus do Araguaia, Barra do Garças/MT. Atualmente é mestranda pelo Programa de pós-graduação em Imunologia Básica e Aplicada na FMRP-USP, onde desenvolve o trabalho de Mestrado no Laboratório de Imunogenética Molecular.

Rodrigo Salazar da Silva é bacharel em Ciências Biológicas pelo Instituto de Biociências da Universidade de São Paulo (IB-USP). Atualmente mestrando do Programa de Pós-graduação em Genética no Departamento de Genética e Biologia Evolutiva na mesma instituição.

Tereza Cristina Leal Balbino é pesquisadora titular do Departamento de Microbiologia do Centro de Pesquisas Aggeu Magalhães – Fiocruz, Recife, PE. Bióloga, com Mestrado e Doutorado na área de Genética pela Universidade Federal de Pernambuco (UFPE).

Thais Campos de Oliveira é mestranda em Biologia Celular e Molecular pela Universidade Federal do Rio Grande do Sul (UFRGS) e farmacêutica pela Faculdade de Ciências Farmacêuticas e Bioquímicas Oswaldo Cruz (FOC).

Tiago Alves Jorge de Souza é bacharel em Ciências Biológicas pelo IBILCE/UNESP (2010). Mestrado pelo Programa de Pós-graduação em Ciências Biológicas/Genética pela FMRP/USP. Atualmente doutorando do mesmo programa.

Tiago Campos Pereira é docente de Genética da Universidade de São Paulo (USP) em Ribeirão Preto. Biólogo pela UNICAMP, com Mestrado e Doutorado em Genética e Biologia Molecular pela mesma instituição. Realizou seu Pós-doutorado em Biotecnologia pela Universidade de Cambridge, Inglaterra.

Lista de símbolos

~	aproximadamente
:	fusão (Cas9 fusionada à GFP → Cas9:GFP)
μg	micrograma
3′	extremidade três prima em ácidos nucleicos
5′	extremidade cinco prima em ácidos nucleicos
A	adenina
C	citossina
F ₀	geração zero
F ₁	primeira geração
F ₂	segunda geração
G	guanina
mL	mililitro
N	qualquer base nitrogenada
ng	nanograma
P ₀	geração parental
R	adenina ou guanina
T	timina
T ₆	seis timinas em sequência
U6	promotor U6 da RNA polimerase III
Y	citossina ou timina
Δ	deleção

Lista de siglas, abreviaturas e acrônimos

aa: aminoácido(s).

ALS (*Amyotrophic Lateral Sclerosis*): esclerose lateral amiotrófica.

BDO: (R,R)-butano-2,3-diol.

BLESS (*direct in situ breaks labelling, enrichment on streptavidin and next-generation sequencing*): marcação direta *in situ* de quebras de dupla fita, enriquecimento com estreptavidina e sequenciamento de última geração.

Cas (*CRISPR associated*): associado a CRISPR.

Cas9 (*CRISPR associated protein 9*): proteína nove associada a CRISPR.

Cas9 HNH⁺: Cas9 nickase, na qual apenas o domínio HNH é cataliticamente ativo (RuvC foi inativado).

Cas9 RuvC⁺: Cas9 nickase, na qual apenas o domínio RuvC é cataliticamente ativo (HNH foi inativado).

Cas9n (*Cas9 nickase*): Cas9 nickase.

Cas9^{wt} (*Cas9 Wild Type*): cas9 tipo selvagem.

cdNA (*complementary deoxyribonucleic acid*): DNA complementar.

ChIP-seq (*Chromatin ImmunoPrecipitation and high-throughput sequencing*): sequenciamento de larga escala de cromatina imunoprecipitada.

Cpf1 (*CRISPR from Prevotella and Francisella 1*): nuclease do sistema CRISPR de *Prevotella* sp. e *Francisella novicida* 1.

CrEdit (*CRISPR-Cas9 mediated genome Editing*): edição genômica mediada por CRISPR.

CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*): repetições palindrômicas pequenas regularmente espaçadas e agrupadas.

CRISPRa (*CRISPR activation*): ativação transcricional via CRISPR.

CRISPRi (*CRISPR interference*): interferência transcricional via CRISPR.

CRISPRm (*multiplex CRISPR*): CRISPR multiplex.

crRNA (*CRISPR-derived RNA*): RNA derivado do loco CRISPR.

dCas9 (*dead Cas9*): Cas9 morta (i.e., cataliticamente inativa).

DECKO (*Double Excision CRISPR Knock Out*): nocaute por CRISPR via dupla excisão.

DG: deleção gênica.

Digenome-seq (*in vitro Cas9-digested whole-genome sequencing*): sequenciamento de genomas completos digeridos por Cas9 *in vitro*.

DmCas9: linhagem transgênica de drosófila expressando Cas9.

DNA (*deoxyribonucleic acid*): ácido desoxirribonucleico.

DrCas9: linhagem transgênica de *Dario rerio* expressando Cas9.

DSB (*double-strand break*): quebra de dupla fita.

dsDNA (*double-stranded deoxyribonucleic acid*): DNA de fita dupla.

dsOligo (*double-stranded oligonucleotide*): oligonucleotídeo de dupla fita.

dsRNA (*double-stranded ribonucleic acid*): RNA de fita dupla.

e.g. (latim, *exempli gratia*): por exemplo.

ELISA (*Enzyme-Linked Immunosorbent Assay*): ensaio de imunoabsorção enzimática.

EMC (*Enzymatic Mismatch Cleavage*): ensaio enzimático de clivagem de bases não pareadas.

ER: enzima de restrição.

ESC (*Embryonic Stem Cell*): célula-tronco embrionária.

eSpCas9 (*enhanced specificity Streptococcus pyogenes Cas9*): Cas9 de alta especificidade.

FACS (*Fluorescence-Activated Cell Sorting*): sorteamento de células ativado por fluorescência.

FISH (*Fluorescent in situ hybridization*): hibridização fluorescente *in situ*.

fCas9 (*FokI fused to Cas9*): Cas9 fusionada a nuclease FokI.

FnCas9 (*Francisella novicida Cas9*): Cas9 de *Francisella novicida*.

GC: guanina e citosina.

GDI: gene de interesse.

GFP (*Green Fluorescent Protein*): proteína verde fluorescente.

GG: dinucleotídeo de guaninas.

GM (*Genetically Modified*): geneticamente modificado.

GMO (*Genetically Modified Organism*): organismo geneticamente modificado.

gRNA (*guide ribonucleic acid*): RNA-guia. O mesmo que sgRNA.

GUIDE-seq (*genome-wide, unbiased identification of double-strand breaks enabled by sequencing*): identificação não enviesada, em escala genômica, de quebras de dupla fita possibilitada por sequenciamento.

GWAS (*Genome Wide Association Study*): estudo de associação em escala genômica.

hCas9 (*human codon-optimized Cas9*): Cas9 otimizada para códons humanos.

HDR (*Homology-Directed Repair*): reparo direcionado por homologia.

HDV (*Hepatitis Delta Virus*): vírus delta da hepatite.

HH (*hammerhead*): ribozima do tipo cabeça de martelo.

HIV (*Human Immunodeficiency Virus*): vírus da imunodeficiência humana.

HRM (*High Resolution Melting*): fusão de alta resolução.

HsCas9 (*Homo sapiens codon-optimized Cas9*): Cas9 otimizada para códons humanos.

HTGTS (*High-Throughput, Genome-Wide, Translocation Sequencing*): sequenciamento de translocações em nível genômico e em larga escala.

i.e. (latim, *id est*): isto é.

IDAA (*Indel Detection by Amplicon Analysis*): detecção de deleções por análise de amplicon.

IDLV (*Integrase-Defective Lentiviral Vector*): vetor lentiviral defectivo para integrase.

Indel (*Insertions and/or deletions*): inserções e/ou deleções.

iPSC (*induced-Pluripotent Stem Cell*): célula-tronco pluripotente induzida.

Kb (*kilobases*): mil bases.

KI: *knock-in*.

KO (*knockout*): noucaute.

kpb: mil pares de bases.

LIC (*Ligation-Independent Cloning*): clonagem independente de ligação.

lncRNA (*long noncoding ribonucleic acid*): RNA longo não codificante.

LPS: lipopolissacarídeo.

LTR (*Long Terminal Repeat*): repetição terminal longa.

MAMA-PCR (*Mismatch Amplification Mutation Assay-Polymerase Chain Reaction*): ensaio de mutação baseado na amplificação por mal pareamento via reação em cadeia da polimerase.

MCR (*Mutagenic Chain Reaction*): reação mutagênica em cadeia.

MIMIVIRE (*Mimivirus Virophage Resistance Element*): elemento do mimivírus para resistência a virófangos.

MIT (*Massachusetts Institute of Technology*): Instituto de Tecnologia de Massachusetts.

MOI (*multiplicity of infection*): multiplicidade de infecção.

Mpb: milhões de pares de bases.

mRNA (*messenger RNA*): RNA mensageiro.

Nat^R (*nourseothricin resistant*): resistente à nourseotricina.

NCBI (*National Center for Biotechnology Information*): Centro Nacional para Informação de Biotecnologia.

Neo^R (*Neomycin resistant*): resistente à neomicina.

NGG (*Nucleotide Guanine Guanine*): qualquer nucleotídeo seguido por duas guaninas.

NHEJ (*Non-Homologous End-Joining*): junção de extremidades não homólogas.

NLS (*Nuclear Localization Signal*): sinal de localização nuclear.

NmCas9 (*Neisseria meningitidis Cas9*): Cas9 endógena de *Neisseria meningitidis*.

nptII (*neomycin phosphotransferase II enzyme*) enzima neomicina fosfotransferase II, que confere resistência à neomicina e kanamicina.

NSD: Nuclease Sítio-Direcionada.

nt: nucleotídeo(s).

ODM (*Oligonucleotide Directed Mutagenesis*): mutagênese direcionada por oligonucleotídeo.

OGM: Organismo Geneticamente Modificado.

paCas9 (*photoactivatable Cas9*): Cas9 fotoativável.

PAM (*Protospacer Adjacent Motif*): motivo adjacente ao protoespaçador.

PAMmer (*PAM-presenting oligonucleotide*): oligonucleotídeo apresentador de PAM.

pb: par(es) de bases.

PCR (*Polymerase Chain Reaction*): reação em cadeia da polimerase.

Pol III: (*RNA Polymerase III*): RNA polimerase III.

RBS (*Ribosome Binding Site*): sítio de ligação de ribossomo.

RCas9 (RNA-targeting Cas9): Cas9 para alvejamento de RNA.

RFLP (Restriction Fragment Length Polymorphism): polimorfismo de comprimento do fragmento de restrição.

RGEn (RNA-guided endonuclease): endonuclease guiada por RNA.

RGN RNP (RNA-guided nuclease ribonucleoprotein): ribonucleoproteína formada por uma nuclease e seu RNA-guia.

RISC (RNA-Induced Silencing Complex): complexo de silenciamento induzido por RNA.

RNA (ribonucleic acid): ácido ribonucleico.

RNAi (RNA interference): interferência por RNA.

RNAP III: (RNA Polymerase III): RNA polimerase III.

RNA-seq (ribonucleic acid sequencing): sequenciamento a partir de transcritos.

RNP (ribonucleoprotein): ribonucleoproteína.

RNP Cas9 (ribonucleoprotein Cas9): ribonucleoproteína Cas9.

RP: Retinite Pigmentosa.

RT (Reverse Transcriptase): transcriptase reversa.

RT:Cas1: Transcriptase Reversa fusionada à Cas1.

RT-qPCR (Reverse Transcription - quantitative Polymerase Chain Reaction): transcrição reversa seguida pela reação em cadeia da polimerase de natureza quantitativa.

RVD (Repeat Variable Diresidues): dirresíduos variáveis repetidos.

SA: Substituição Alélica.

SaCas9 (Staphylococcus aureus Cas9): Cas9 endógena de *Staphylococcus aureus*.

scFv (single-chain variable fragment): fragmento variável de cadeia simples.

SG: Sequência-Guia.

sgRNA (single guide ribonucleic acid): RNA-guia de fita única. O mesmo que gRNA.

shRNA (short hairpin RNA): pequeno RNA de interferência em forma de grampo.

siRNA (small interfering RNA): pequeno RNA de interferência.

snoRNA (small nucleolar ribonucleic acid): pequeno RNA do nucléolo.

SNP (Single Nucleotide Polymorphism): polimorfismo de um único nucleotídeo.

snRNA (small nuclear ribonucleic acid): pequeno RNA do núcleo.

SpCas9 (Streptococcus pyogenes Cas9): Cas9 endógena de *Streptococcus pyogenes*.

SpCas9-HF1 (Streptococcus pyogenes Cas9 High Fidelity 1): Cas9 de alta fidelidade de *Streptococcus pyogenes*.

ssDNA (single-stranded deoxyribonucleic acid): DNA de fita única.

ssODN (single-stranded oligodeoxyribonucleotide): oligodesoxinucleotídeo de fita única.

ssRNA (single-stranded ribonucleic acid): RNA de fita única.

St1Cas9 (Streptococcus thermophilus Cas9): Cas9 endógena de *Streptococcus thermophilus*.

TALE (Transcription Activator-Like Effector): efetor semelhante ao ativador de transcrição.

TALEN (Transcription Activator-Like Effector Nuclease): nuclease fusionada ao efetor semelhante ao ativador de transcrição.

TNF (*Tumor Necrosis Factor*): fator de necrose tumoral.

tracrRNA (*trans-activating CRISPR-derived RNA*): RNA transativador derivado do loco CRISPR.

tRNA (*transfer ribonucleic acid*): RNA transportador.

TSS (*Transcription Start Site*): sítio de início da transcrição.

USDA (*United States Department of Agriculture*): Departamento de Agricultura dos Estados Unidos.

USPTO (*United States Patent e Trademark Office*): Escritório de Marcas e Patentes dos Estados Unidos.

UTR (*Untranslated Region*): região não traduzida.

WES (*Whole Exome Sequencing*): sequenciamento do exoma completo.

WGS (*Whole Genome Sequencing*): sequenciamento do genoma completo.

WT (*Wild Type*): tipo selvagem.

ZFN (*Zinc Finger Nuclease*): nuclease dedo de zinco.

Sumário

	Prefácio	25
	Tiago Campos Pereira	
	Prefácio especial, por um dos pioneiros de CRISPR	27
	Luciano A. Marraffini	

Capítulo 1.	Introdução	29
	Tereza Cristina Leal Balbino, Maria Paloma Silva de Barros, Arthur Casulli de Oliveira, Danilo Pinhal, Bruno Henrique Garcia e Tiago Campos Pereira	
	1. Histórico da descoberta do loco CRISPR	30
	2. Função biológica de CRISPR-Cas	31
	3. Diferenças entre sistemas I, II e III	31
	4. Origem e evolução dos sistemas CRISPR-Cas	33
	5. CRISPR em número de publicações	35
	6. Termos e definições gerais	36
	7. Conclusões	37

Capítulo 2.	Visão geral do mecanismo básico de ação	39
	Gabriel Viliod Vieira, Nerry Tatiana Cecílio, Letícia Magalhães Arruda e Katiuchia Uzzun Sales	
	1. Visão geral sobre moléculas envolvidas e mecanismos básicos de ação	40
	2. Moléculas envolvidas na técnica de CRISPR	40
	2.1 A nuclease Cas9	40
	2.2 RNA-guia (<i>Single guide</i> RNA)	40
	2.3 DNA-alvo	42
	3. Mecanismo de ação de Cas9/sgRNA	43
	3.1 Clivagem de corte duplo	43
	3.2 Clivagem de corte simples	44
	4. Reparo do DNA: NHEJ e HDR	44
	4.1 Junção de Pontas Não Homólogas (NHEJ)	44
	4.2 Reparo Direcionado por Homologia (HDR)	45
	5. Conclusão	47
	Apêndices	47

Capítulo 3.	CRISPR <i>toolbox</i>: a versatilidade da técnica	51
	Caroline Junqueira, Patrícia Renck Nunes, Gabriela Molinari Roberto e María Sol Brassesco Annichini	

1. Introdução	52
2. Edição de DNA	52
2.1 Inativação gênica	52
2.2 Integração de transgenes	54
2.3 Substituição alélica	55
2.4 Deleção de genes inteiros	55
3. Marcação de DNA	55
4. Regulação da expressão gênica	58
4.1 Para indução da expressão gênica	58
4.2 Para repressão da expressão gênica	61
5. Clivagem de RNAs	63
6. Mapeamento de genes	64
7. Rastreamento de RNAs	67
8. Conclusões	69

Capítulo 4.	Aplicações da técnica	71
	Geraldo Aleixo Passos, Cesar Augusto Speck Hernandez, Larissa Cotrim de Sousa, Rafaela de Freitas Martins Felício e Tiago Alves Jorge de Souza	

1. Introdução	72
2. Aplicações na pesquisa básica	72
2.1 Identificação da função gênica	72
2.2 Construção de modelos para pesquisa (células e organismos)	73
3. Aplicações na medicina	75
3.1 Combate a patógenos	75
3.2 Combate ao câncer	76
3.3 Terapia genética	77
3.4 Controle de vetores de doenças	77
4. Aplicações na agricultura e veterinária	78
4.1 Combate a pragas agrícolas	78
4.2 Geração de variação genética – novos cultivares não transgênicos	78
4.3 Animais resistentes a doenças	79
5. Aplicações na indústria	79
5.1 Produção de biomoléculas	79
5.2 Produção de biocombustíveis	80
5.3 Produção de biomateriais	80
6. Conclusões	81

Capítulo 5.	Comparação com outras técnicas para <i>knockout</i> e <i>knockdown</i>: RNAi, Gene Targeting, ZFNs e TALENs	85
	Leonardo Chicaybam, Bárbara Peixoto e Martin Bonamino	

1. Introdução	86
2. Interferência por RNA (RNAi)	86
3. <i>Gene Targeting</i> e Cre/loxP	88
4. Nucleases Dedo de Zinco (ZFNs)	89
5. <i>Transcription activator-like effector nucleases</i> (TALENs)	91
6. Conclusões	93

Capítulo 6.	Variantes naturais e artificiais de Cas9 e sgRNAs	97
	Julio Cesar Cetrulo Lorenzi	

1. Introdução	98
2. Cas9 clássica de <i>Streptococcus pyogenes</i> (SpCas9)	98
3. <i>Nickases</i>	98
4. Cas9 com ambos os sítios catalíticos inativos (<i>dead</i> Cas9 - dCas9)	100
5. Entendimento e desenvolvimento da região PAM.	101
6. Cas9 de alta precisão	102
7. Sistemas de ativação indireta da Cas9	103
8. Modificações nos sgRNAs	105
9. Cpf1	105
10. Conclusões	106

Capítulo 7.	Obtenção de Cas9 e gRNAs	109
	Rodrigo Salazar da Silva, Erika Kague, Tiago Campos Pereira, Natalia Nardelli Gonçalves e Carlos Eduardo Ambrósio	

1. Introdução	110
2. A nuclease Cas9	110
2.1 Formatos disponíveis: vetores, mRNAs e proteínas.	110
2.2 Vias de obtenção: acadêmicas, comerciais e síntese	112
3. O gRNA	117
3.1 Formatos disponíveis: vetores e RNAs	117
3.2 Vias de obtenção: acadêmicas, comerciais e síntese	117
4. Busca por gRNAs indisponíveis comercialmente	119
5. Desenho de gRNAs	119
5.1 Posição do gRNA em relação ao gene	119
5.2 Outros aspectos	121
6. Construção de bibliotecas de gRNAs	123
6.1 <i>Pooled libraries</i> – visão geral	124
6.2 Tipos de <i>pooled libraries</i>	128
7. Conclusões	130

Capítulo 8.	Controles experimentais, formas de análise e cuidados especiais.	135
	Bruno Trevenzoli Favero, Oliver Kemp e Tiago Campos Pereira	
1.	Introdução	136
2.	Controles experimentais.	136
2.1	Controles negativos	136
2.2	Controles positivos.	136
2.3	De especificidade.	136
3.	Confirmação da edição	138
3.1	Análises de DNA.	139
3.2	Análises de RNA.	141
3.3	Análises de Proteína	142
4.	Cuidados especiais.	142
4.1	Técnicas para identificação de <i>off-targets</i>	142
4.2	Efeito mosaico	146
5.	Conclusões	148
Capítulo 9.	Métodos para o uso de CRISPR	151
	Marcelo Bassalo, Gleidson Silva Teixeira, Gonçalo Amarante Guimarães Pereira, Andréia Prata Vieira, Marie Anne Van Sluys, Joan Barau, Kaio Cesar Simiano Tavares, Juliana Borsoi e Lygia da Veiga Pereira	
1.	Introdução	152
2.	Bactérias	152
2.1	Introdução	152
2.2	Regras para desenho e construção de gRNAs e DNA doador para bactérias	152
2.3	Edição da sequência nucleotídica	155
2.4	Nocautamento e deleções gênicas	157
2.5	Inserção de genes heterólogos	158
3.	Leveduras	159
3.1	Introdução	159
3.2	Nocautamento de genes	160
3.3	Integração de transgenes	161
3.4	Integração de vias metabólicas	162
4.	Fungos filamentosos	163
4.1	Introdução	163
4.2	Exemplos.	163
5.	Plantas	165
5.1	Introdução	165
5.2	Inserção de transgene (KI) em <i>Arabidopsis</i>	166
5.3	Deleção de sequências genômicas em arroz	168
5.4	Geração de variabilidade genética livre de DNAs exógenos	169
5.5	Regulação da expressão gênica <i>in planta</i>	171
5.6	Próximos passos do uso de CRISPR em plantas	171
6.	Animais	173
6.1	<i>Drosophila melanogaster</i> (Chen et al., 2016)	173

6.2	<i>Caenorhabditis elegans</i> (Paix et al., 2014)	173
6.3	<i>Xenopus laevis</i> (Wang et al., 2015)	176
6.4	<i>Zebrafish</i> (Di Donato et al., 2016)	176
6.5	Camundongos (Maddalo et al., 2014)	177
6.6	Ratos (Bakondi et al., 2016)	177
6.7	Cabras (Ni et al., 2014)	178
6.8	Bovinos (Heo et al., 2015)	178
6.9	Cães (Zou et al., 2015)	179
6.10	Suínos (Yang et al., 2015)	179
6.11	Macacos (Chen et al., 2015; Kang et al., 2015)	180
6.12	Conclusões	180
7.	Células humanas	181
7.1	Introdução	181
7.2	Substituição alélica	182
7.3	Deleção de sequências gênicas.	183
7.4	Inativação gênica	184
7.5	Integração de sequências exógenas (<i>knock-in</i>)	185
7.6	Edição genética em espermatogônias	187

Capítulo 10. **Otimizações da técnica**193

Tiago Alves Jorge de Souza, Rafael Sanguinetti Czepielewski,
Daiana Renck, Cristina Bonorino, Jayme Augusto de Souza-Neto
e Tiago Campos Pereira

1.	Introdução	194
2.	Inibição da via NHEJ	194
3.	<i>Multiplex</i> CRISPR	197
4.	Reação Mutagênica em Cadeia (MCR)	202
5.	Espécies transgênicas para Cas9	209
6.	Conclusões do capítulo	212

Capítulo 11. **Questões éticas, legais, ambientais e de pioneirismo**215

Nilza Maria Diniz, Andréa Almeida Carneiro, Newton Portilho Carneiro,
Camila Tenorio França, Maria Betânia Melo de Oliveira
e Giulliana Augusta Rangel Gonçalves

1.	Introdução	216
2.	Éticas	216
2.1	Bioética e genética	216
2.2	A edição de embriões humanos	217
2.3	A questão moral	220
3.	Legais	220
3.1	A regulamentação da edição genética	220
3.2	O quadro regulatório mundial	221
4.	Ambientais	224
4.1	Introdução	224
4.2	Potenciais aplicações da técnica MCR	226
4.3	Cuidados no uso da técnica e regulamentações vigentes	227

4.4	Riscos ambientais da dispersão de genes derivados de MCR	227
4.5	Controvérsias sobre a técnica MCR e sua segurança	228
4.6	Conclusão	229
5.	Pioneirismo	229
5.1	As patentes e o mercado da biotecnologia	229
5.2	A disputa Doudna-Charpentier versus Zhang	230
6.	Conclusões	234

Capítulo 12.	Tópicos de fronteira	237
	Luciana Rodrigues Carvalho Barros, Thais Campos de Oliveira, Augusto Schrank e Tiago Campos Pereira	
1.	Introdução	238
2.	CRISPR <i>goes retro</i>	238
3.	MIMIVIRE: CRISPR de vírus gigantes?	239
3.1	Mimivírus	239
3.2	O mecanismo MIMIVIRE.	240
4.	Tópicos especiais sobre CRISPR	241
5.	Conclusões	242

Glossário	245
----------------------------	-----

Índice remissivo	249
-----------------------------------	-----

Prefácio

Tiago Campos Pereira

Regularmente presenciamos o surgimento de técnicas revolucionárias na Genética. Assim foi com a PCR em 1983, a Interferência por RNA (RNAi) em 1998 e agora a CRISPR em 2012/13.

Um detalhe inesperado entre RNAi e CRISPR é que elas parecem ser *versões espelhadas de um mesmo processo*. A RNAi é o sistema *eucariótico* de defesa antiviral, baseado em uma endonuclease guiada por RNA. Ela promove o silenciamento gênico pós-transcricional por meio da clivagem do RNA-alvo. Por sua vez, CRISPR é um sistema *procariótico* de defesa antiviral, baseado em endonuclease guiada por RNA, capaz de inativar genes por meio da clivagem do DNA-alvo. Contudo, apesar de a RNAi não ser tipicamente aplicável em procariotos, a CRISPR é exequível em eucariotos.

Um terceiro elemento foi recentemente agregado a este cenário, tornando-o ainda mais inusitado: MIMIVIRE – o *sistema antiviral de vírus* (vide cap. 12). Semelhanças entre RNAi (eucariotos), CRISPR (procariotos) e MIMIVIRE (vírus) levam alguns a conjecturar que uma nova técnica de edição genética surja nos próximos anos a partir desta última, completando o quadro.

Diante da avalanche em escala global de artigos utilizando CRISPR, a elaboração de um livro que pudesse auxiliar o pesquisador quanto a este tema tornou-se inexorável. Assim, esta obra foi planejada objetivando apresentar o vasto tema “CRISPR” de maneira clara, didática, abrangente e atualizada, para que fosse acessível tanto a graduandos como a docentes. Os primeiros dois capítulos abordam aspectos fundamentais: o histórico da descoberta do loco CRISPR, seu papel na defesa antiviral em bactérias, a visão geral do mecanismo básico de ação da técnica e as moléculas envolvidas no processo.

Os três capítulos seguintes (caps. 3 a 5) discorrem sobre a imensa versatilidade da técnica, atuando no DNA e RNA por meio de interações distintas da Cas9 (ligação ou clivagem), as aplicações de CRISPR em diversas áreas e a comparação desta técnica com outras de natureza semelhante. Os dois capítulos posteriores (caps. 6 e 7) discutem sobre as variantes naturais e artificiais de Cas9 e do RNA-guia (os dois pilares da técnica de CRISPR) e como obtê-los para realização dos estudos.

Os capítulos 8 a 10 abordam questões práticas da técnica: controles necessários para condução do experimento, formas para confirmar a edição genética, exemplos de métodos em diversos modelos (de bactérias a células humanas) e estratégias para otimização dos procedimentos. Por fim, os últimos dois capítulos (11 e 12) apresentam questões controversas sobre a técnica (aspectos legais, bioética, biossegurança e pioneirismo) e alguns tópicos de fronteira na área.

Esta obra é o terceiro volume de uma série chamada “**Introdução a ...**”, por mim organizada, que objetiva apresentar de maneira clara as técnicas e tópicos recentes da genética e biologia molecular aos alunos e pesquisadores brasileiros. Os dois primeiros volumes – **Introdução à técnica de Interferência por RNA - RNAi** (2013) e **Introdução**

ao mundo dos microRNAs (2015) – têm sido um grande sucesso, estando disponíveis na livraria virtual da Sociedade Brasileira de Genética.

Pretendendo elaborar uma obra abrangente sobre um tema tão recente e importante, uma nova rede de colaboradores foi nucleada neste terceiro volume ao longo de sete meses, para que ele fosse amplo (12 capítulos), igualmente didático (~ 50 figuras), com diferentes pontos de vista (51 colaboradores), visando proporcionar ao leitor um texto estruturalmente organizado, coeso, extensivo e de fácil leitura.

Registro aqui meus agradecimentos a todos que auxiliaram direta e indiretamente na produção desta obra (autores, colaboradores, revisores, editores e parceiros), às agências de fomento CNPq, CAPES, FAPESP e Fundações de Amparo à Pesquisa dos estados de todos os autores, por fornecerem todo o arcabouço necessário para a elaboração deste livro.

Em especial, agradeço aos meus pais, por me fazerem enxergar além do que meus olhos podem ver.

Prefácio especial, por um dos pioneiros de CRISPR

Luciano A. Marraffini

*Associate Professor and Head of the Laboratory of Bacteriology
The Rockefeller University
New York, USA*

If you are reading this it is most likely that you are interested in changing the genetic code of some creature. You are not the only one: this has been a humankind goal since the beginning of time. Early on, breeding was used to introduce modifications into organisms. More recently, mutagens were used to generate genetic diversity and develop new valuable properties in animals and plants. The discovery of DNA as the cell genetic material led to the concept of genetic engineering, where new functions could be introduced by the rational modification of the organism DNA sequence. In the '70s, the recombinant DNA technology made possible the efficient genetic engineering of bacteria. However, similar tools for the manipulation of higher eukaryotes were not available.

Enter CRISPRs. In essence, CRISPRs are arrays of repetitive DNA sequences, hence the acronym: Clustered Regularly Interspaced Short Palindromic Repeats. CRISPR loci and their associated Cas proteins constitute an adaptive immune system of bacteria and archaea. They produce short RNAs, known as CRISPR RNAs, that guide Cas nucleases to their target sequence in the genome of viruses and plasmids that infect bacteria. The RNA-guided DNA cleavage of Cas nucleases destroys the invader, providing immunity to the prokaryotic host. A fundamental feature of this immune response is that bacteria and archaea can re-program their CRISPR systems to produce a new CRISPR RNA and target a new invader. Understanding how CRISPR systems can be reprogramed by prokaryotes has allowed us to program Cas nucleases to cleave any sequence we want and thus to find the perfect tool for the genetic engineering of higher eukaryotes. This book provides a comprehensive overview of the technique, providing a superb guide on how to use Cas9 for genome engineering, how to troubleshoot your experiments, and how to use other applications originated from the CRISPR toolbox. It should be an essential manual in every lab that uses the CRISPR-Cas9 technology.

Genetic engineering with CRISPR-Cas can be so precise that it is also referred as “genome editing”, where we can edit the DNA sequence of any organism with as much precision as we edit a text with a word processing program. In addition, the technique is so easy to carry out and so efficient that it has spread to labs around the globe with the speed of light leading to the “CRISPR revolution.”

In truth, the CRISPR revolution started much earlier, with bioinformatics works that predicted the function of these loci as anti-phage and anti-plasmid systems and the subsequent experimental demonstration of these predictions. I was involved in this early experimental work. Still as a student in the laboratory of Olaf Schneewind I became fascinated with CRISPR systems and started the work that showed their anti-plasmid function in bacteria. I continued my work in the lab of Erik Sontheimer, where we demonstrated that CRISPR

immunity required the destruction of the invading DNA and not RNA as it was predicted by the preceding bioinformatics analyses. This was a key finding for the future applications of CRISPR in genome editing (it would not be possible to introduce mutations in human cells with an RNA-cleaving system), which we highlighted in 2008 when we published our findings. Later on, these findings were put to work for genetic engineering when I teamed up with Feng Zhang to introduce mutations in human cells using CRISPR immunity.

The journey from these first discoveries to the current widespread use of the technology has been very rewarding for me personally. But more importantly, this journey exemplifies the tremendous value of basic research. Like the discovery of restriction enzymes more than forty years ago, the discovery of CRISPR resulted from scientific endeavors aimed at understanding how prokaryotes interact with the mobile genetic elements that infect them; an area of biology that a priori does not have any potential applications to human health. And like restriction enzymes, Cas nucleases enabled a revolution in genetics, highlighting the importance of studying interesting problems, regardless of their potential for biotechnology.

Introdução

Capítulo

1

Tereza Cristina Leal Balbino^{1,a}, Maria Paloma Silva de Barros^{1,a},
Arthur Casulli de Oliveira^{2,b}, Danillo Pinhal^{2,b}, Bruno Henrique Garcia^{3,c}
e Tiago Campos Pereira^{3,4,c,d}

¹*Dept.º de Microbiologia, CPqAM/FIOCRUZ, PE - Brasil*

²*Dept.º de Genética, Instituto de Biociências, UNESP, SP - Brasil*

³*Dept.º de Biologia, FFCLRP - USP, SP - Brasil*

⁴*Programa de Pós-Graduação em Genética, FMRP - USP, SP - Brasil*

^a*Autoria das seções: 1 a 3*

^b*Autoria da seção: 4*

^c*Autoria da seção: 5*

^d*Autoria das seções: 6 e 7*

Estruturação do capítulo

1. Histórico da descoberta do loco CRISPR
2. Função biológica de CRISPR-Cas
3. Diferenças entre sistemas I, II e III
4. Origem e evolução dos sistemas CRISPR-Cas
5. CRISPR em número de publicações
6. Termos e definições gerais
7. Conclusões

1. Histórico da descoberta do loco CRISPR

Em 1987, o pesquisador Yoshizumi Ishino e colaboradores da Universidade de Osaka (Japão) identificaram um loco (região) peculiar no genoma da bactéria *Escherichia coli*, constituído por uma configuração incomum: *sequências repetidas* e *sequências espaçadoras* intercaladas e de função desconhecida (figura 1). Tais sequências foram investigadas independentemente em 1993, e em 2000 elas foram identificadas nos genomas de diferentes bactérias e Archaea (Mojica et al., 1993; Mojica et al., 2000). Somente no ano de 2002, a sigla CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats* ou Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespaçadas) foi criada para denominar tais sequências (Jansen et al., 2002).

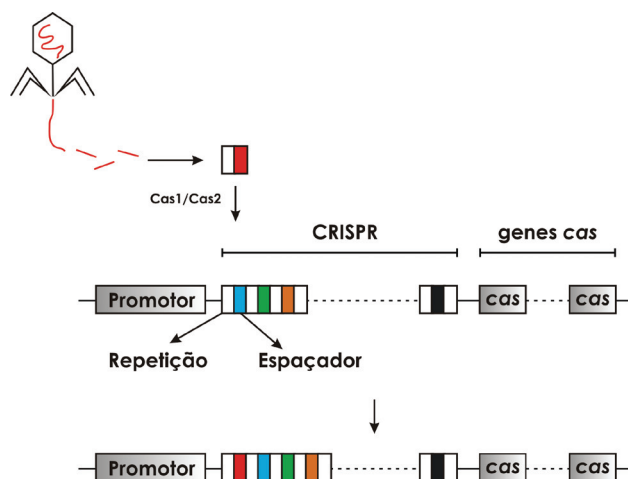


Figura 1. Sistema de defesa CRISPR-Cas. Loco CRISPR e etapa de adaptação. As proteínas Cas integram o DNA exógeno (viral ou plasmidial) como novo espaçador no loco CRISPR. Imagem: Pereira, TC (2016) baseado em Jiang e Marraffini (2015).

Ao longo dos anos, um conjunto de genes fisicamente muito próximo ao loco CRISPR foi identificado e nomeado de genes *Cas* (*CRISPR associated genes*) (Jansen et al., 2002), exemplificados por nucleases, polimerases e helicases que viriam a se mostrar futuramente como elementos centrais no funcionamento do loco como um todo.

Em 2005, três grupos independentes de pesquisadores mostraram que as sequências espaçadoras (ou simplesmente *espaçadores*) têm origem extracromossômica, isto é, são derivadas de plasmídeos ou de vírus. Adicionalmente, foi também descrito que vírus são incapazes de infectar com sucesso bactérias que possuem espaçadores cujas sequências são correspondentes a trechos de seus genomas (Bolotin et al., 2005).

Baseado nesses achados, em 2005 foi postulada a hipótese de que CRISPR-Cas seria um *sistema imune adaptativo* de procariotos, no qual os espaçadores serviriam como “memória de invasões anteriores” (Mojica et al., 2015). Assim, moléculas de RNA geradas a partir desses espaçadores seriam complementares ao patógeno (re)invasor, permitindo combatê-lo de maneira sequência-específica. Uma série de estudos subsequentes revelou ser esta hipótese verdadeira.

2. Função biológica de CRISPR-Cas

O sistema CRISPR-Cas é uma maquinaria adaptativa de defesa que permite aos procariotos protegerem-se contra a (re)invasão de elementos genéticos móveis indesejados, tais como bacteriófagos, transposons e plasmídeos, assemelhando-se a um sistema imune adaptativo nesses organismos (Sampson e Weiss, 2014). O mecanismo de defesa se resume a três etapas: (i) *adaptação*, (ii) *biogênese do crRNA* e (iii) *ação contra o invasor*.

A primeira etapa – adaptação – é marcada pela aquisição de um novo espaçador no loco CRISPR (Makarova et al., 2011). Isto ocorre quando a bactéria é infectada pela primeira vez pelo vírus, por exemplo. Nesse momento, algumas enzimas codificadas pelos genes *Cas* (*Cas1* e *Cas2*) clivam o DNA do patógeno em pequenos segmentos (24-48 pares de bases) e os integram no loco CRISPR como novos espaçadores (ou seja, entre as sequências repetidas). A partir desse momento, a bactéria está imunizada contra futuras invasões deste mesmo agente.

Na segunda fase – biogênese do crRNA (*CRISPR-derived RNA*) – ocorre a transcrição ininterrupta do loco CRISPR. Esse processo é mediado pela *Sequência Líder* (L), uma região com alto conteúdo de Adenina-Timina que serve como sítio promotor. A transcrição deste loco gera um “precursor do RNA de CRISPR” (ou pré-crRNA), contendo várias sequências repetidas e vários espaçadores em um único RNA longo (vide cap. 2, figura 1). Em seguida, esse pré-crRNA é processado, dando origem a diversos RNAs menores – os crRNAs –, cada qual correspondendo a um espaçador distinto (Makarova et al., 2011).

Por fim, no terceiro momento – ação contra o DNA invasor – os crRNAs maduros, junto com as proteínas *Cas*, formam complexos que reconhecem a sequência genética exógena (plasmídeo, transposon ou vírus) e a destroem (figura 2) (Makarova et al., 2011). Este processo em muitos aspectos recapitula o mecanismo de Interferência por RNA (RNAi) observado em eucariotos (Barrangou e Horvath, 2012; Westra et al., 2014).

Este sistema e seus elementos são variáveis entre as diferentes espécies em termos de ocorrência, composição gênica, sequências, número e tamanho (Barrangou e Marraffini, 2014). Em especial, a composição dos espaçadores pode variar até mesmo entre indivíduos da mesma espécie, pois elas derivam de contatos anteriores com diferentes patógenos. Embora o funcionamento do sistema CRISPR-Cas ainda não seja compreendido em sua totalidade, pesquisas têm revelado a importância deste para inúmeras aplicações: resistência bacteriana a fagos; controle da disseminação de genes, via transferência horizontal; genotipagem de cepas (baseada na hipervariabilidade dos espaçadores) e estudo da dinâmica populacional microbiana (Horvath e Barrangou, 2010). Também já existem diversos estudos utilizando esta maquinaria como forma eficaz de engenharia genética, não apenas em microrganismos, mas também em plantas de interesse econômico (Jiang e Marraffini, 2015).

3. Diferenças entre sistemas I, II e III

O sistema CRISPR-Cas procariótico codifica proteínas que atuam como um sistema imune adaptativo contra infecções virais e plasmidiais. A imunidade é mediada pelas nucleases *Cas* e pequenos RNAs-guias (crRNAs), que especificam o sítio de clivagem dentro do genoma invasor. Este sistema é formado por dois componentes principais: as proteínas *Cas*, que funcionam como catalisadores, e o loco CRISPR, como memória genética.

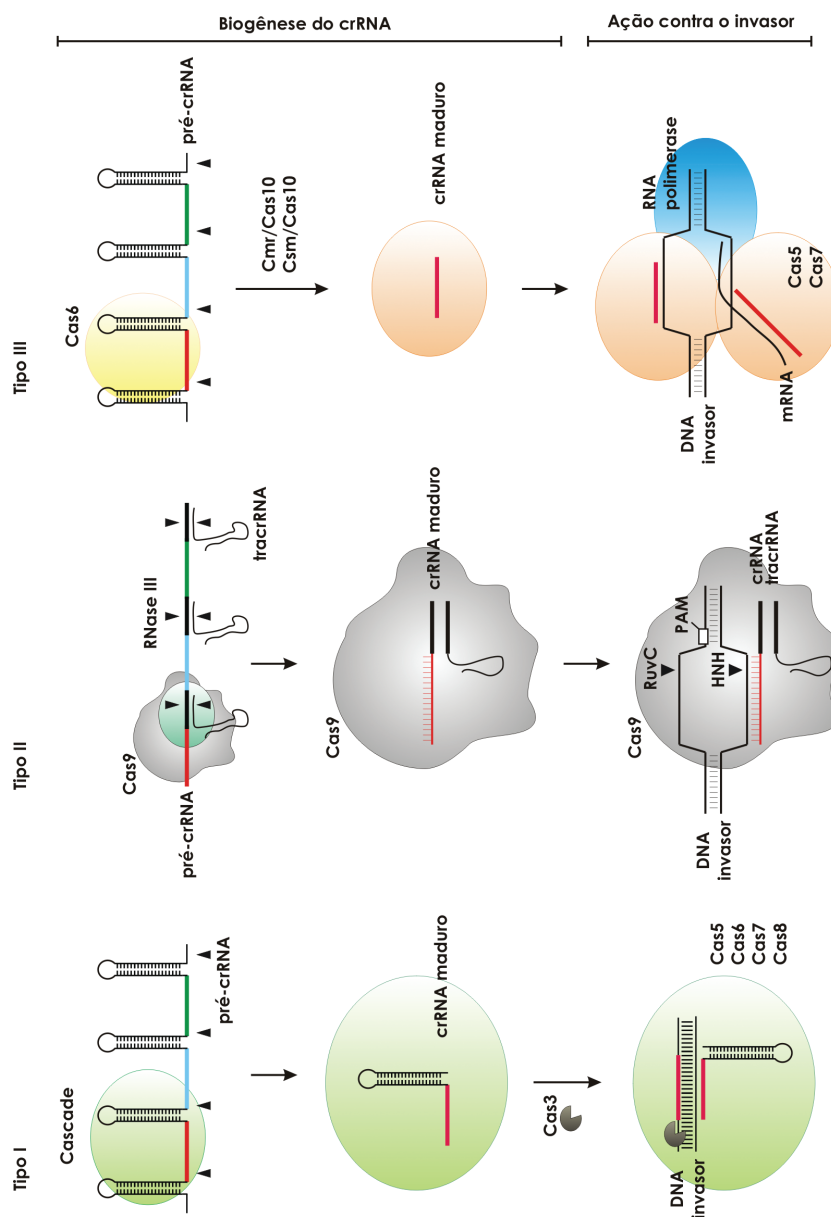


Figura 2. Sistemas CRISPR-Cas Tipos I, II e III. Imagem: Pereira, TC (2016), baseado em Jiang e Marraffini (2015).

Inicialmente foram descritos três tipos diferentes de sistema CRISPR-Cas (Tipo I, Tipo II e Tipo III) e onze subtipos (IA-F, IIA-C e IIIA-B). Nove proteínas Cas fazem parte dos sistemas: Cas1 e Cas2 estão presentes nos três tipos (Tipo I, Tipo II e Tipo III); Cas5, Cas6 e Cas7 nos Tipo I e Tipo III. São consideradas como proteínas exclusivas para cada tipo de sistema: Cas3 e Cas8 (Tipo I), Cas9 (Tipo II) e Cas 10 (Tipo III). Outras proteínas são subtipos e denominadas tipo-dependentes, presentes em determinadas espécies bacterianas, figura 2 (Jiang e Marraffini, 2015).

Algumas variações caracterizam cada um dos três sistemas CRISPR, que são agrupados de acordo com a conservação dos genes *cas* e da organização do operon (Mali et al., 2013).

O sistema CRISPR-Cas Tipo I é definido por um complexo associado a CRISPR para defesa antiviral (cascade) e Cas3. O complexo reúne as Cas: Cas5, Cas6, Cas7 e Cas8. É direcionado ao DNA invasor por meio de reconhecimento do domínio PAM (*Protospacer Adjacent Motif* ou Motivo Adjacente ao Protoespaçador). A Cas3 é responsável pela destruição do material genético exógeno (figura 2).

O sistema CRISPR-Cas Tipo II é o sistema mais simples de todos. É definido pela presença da endonuclease Cas9, guiada pelo crRNA. O gene *cas9* é o único necessário para o combate ao DNA invasor, neste sistema. Durante uma infecção, o complexo formado pela Cas9 destrói o genoma viral, baseando-se também no domínio PAM. A enzima Cas9 comporta dois domínios com atividades nucleásicas (RuvC e HNH), que participam da imunidade (figura 2).

Por fim, o sistema CRISPR-Cas Tipo III parece ser o mais complexo, subdividido em dois subtipos, III-A e III-B. Os dois subtipos estão associados à proteína Cas10, entretanto, podem ser diferenciados pelos genes acessórios: *csm* no subtipo III-A e *cmr* no subtipo III-B. Além disso, eles não requerem domínios PAM para o reconhecimento da sequência-alvo ou para sua clivagem. Diferente dos outros sistemas, o Tipo III reconhece e destrói o RNA ou DNA como material genético exógeno. O complexo tipo III-A degrada o DNA invasor e o Tipo III-B apenas o RNA invasor (figura 2) (Hatoum-Aslan e Marraffini, 2014; Jiang e Marraffini, 2015).

4. Origem e evolução dos sistemas CRISPR-Cas

Ao longo do processo evolutivo, a aquisição de genes responsáveis por promover mecanismos de defesa contra agentes patogênicos tem sido de vital importância para a sobrevivência dos organismos. De modo geral, os organismos procariotos (Bactérias e Árqueas) constantemente recebem por transferência horizontal, via mecanismos de conjugação (plasmídeos), transformação (DNA exógeno) e/ou transdução (via bacteriófagos), material genético exógeno. Entretanto, apenas uma mínima porção desse material genético adquirido confere vantagens adaptativas para esses organismos, o que favoreceu a evolução concomitante de mecanismos de defesa mantenedores da integridade do genoma.

Análises evolutivas têm evidenciado que os mecanismos de defesa componentes dos sistemas imunes tiveram origens independentes entre os domínios da vida, a despeito de sua similaridade de ação (Stern, 2014). Então, como justificar a analogia funcional entre esses mecanismos? A detalhada observação dos sistemas biológicos mostra que existem algumas poucas soluções básicas para os problemas envolvendo parasitas moleculares, como por exemplo: (i) o reconhecimento do alvo (enzimas de restrição, siRNAs, miRNAs, etc.), (ii) a via de apoptose (morte celular programada), evitando a disseminação do parasita e (iii) a memória de contaminações passadas (CRISPR-Cas, piRNA e imunidade adaptativa). Desta maneira, ao longo do processo evolutivo é perfeitamente possível que sistemas imunes funcionalmente análogos surjam de forma independente nos diversos grupos de organismos.

O surgimento do sistema CRISPR-Cas, identificado recentemente como um sistema imune adaptativo nos organismos procariotos, é alvo de diversos debates nos dias de hoje, principalmente devido às diferenças encontradas entre os sistemas de Árqueas e Bactérias. Em Árqueas, os locos CRISPR e Cas são mais complexos e geralmente mais longos do que

nas Bactérias, e ainda possuem mais cópias no genoma (Mojica et al., 2010). Além disso, as sequências repetitivas apresentam baixa homologia entre os genomas de Árqueas e Bactérias (Cooper e Overstreet, 2014).

O sistema CRISPR-Cas encontrado nesses organismos procariotos atua de maneira análoga à Interferência por RNA (RNAi) dos organismos do domínio Eukária. Similarmente aos mecanismos de ação de RNAi conhecidos, CRISPR-Cas evoluiu como um mecanismo de clivagem das sequências-alvo via interação de seu complexo proteico a pequenos RNAs-guia não codificadores (Bhaya et al., 2011). Desse modo, analogias aos mecanismos de ação de *small interfering RNAs* (siRNAs), microRNAs (miRNAs) e *piwi-interacting RNAs* (piRNAs) podem ser traçadas, havendo uma altíssima semelhança entre este último e o sistema CRISPR-Cas. Entretanto, até o presente momento não foram detectadas estruturas gênicas homólogas entre CRISPR-Cas e os sistemas de RNAi dos Eucariotos, sugerindo que esses dois mecanismos surgiram de maneira independente durante a história evolutiva (Shabalina e Colin, 2008). Por outro lado, dados relativos aos mecanismos moleculares envolvendo CRISPR-Cas e principalmente os piRNAs são ainda escassos, e trabalhos futuros poderão detectar sequências homólogas entre esses dois sistemas, o que seria indicativo de uma origem evolutiva comum (Saha e Garret, 2011).

Hoje em dia, aproximadamente 50% das Bactérias já sequenciadas apresentam o sistema CRISPR-Cas em seu genoma, enquanto 90% das Árqueas sequenciadas o contêm (Weinberger et al., 2012). Entretanto, Weinberger et al. (2012) apontam que a maioria das Árqueas já sequenciadas são termofílicas, enquanto grande parte das Bactérias sequenciadas são mesofílicas. Dessa maneira, a presença do sistema CRISPR-Cas parece estar mais associada às condições ambientais, tais como à alta temperatura, do que propriamente ao domínio filogenético. Organismos oriundos de regiões com alta temperatura tendem a apresentar baixa taxa mutagênica, uma vez que é altamente provável que uma mutação acarrete problemas para a sobrevivência nas referidas condições. Assim, a baixa taxa de mutação dos vírus parasitas destes organismos permite que o sistema CRISPR-Cas também se mantenha estável e capaz de combater a infecção viral. De modo inverso, Weinberger et al. (2012) argumentam que em condições mesofílicas a alta taxa mutagênica dos vírus supera a capacidade de adição dos espaçadores no sistema CRISPR-Cas, o que acarretaria na perda da eficiência deste sistema por parte dos organismos mesofílicos.

Por outro lado, utilizando um modelo matemático, Haerter et al. (2011) propõem que a coexistência do sistema CRISPR-Cas e bacteriófagos é possível mesmo quando o número de espécies de fagos supera a capacidade imunológica do sistema. Os autores também argumentam que se o custo metabólico do sistema CRISPR-Cas estiver relacionado ao seu tamanho, poderia haver uma diminuição do número de espaçadores nas Bactérias mesofílicas, o que poderia explicar seu menor tamanho em Bactérias em relação às Árqueas.

Makarova et al. (2011) propõem que o primeiro sistema CRISPR-Cas tenha surgido em Árqueas termofílicas e foi posteriormente adquirido pelas Bactérias por transferência horizontal. Esta proposta foi elaborada tomando-se por base a polimerase do sistema CRISPR, formada pelo gene *cas10*, presente no sistema Tipo III e que, segundo reconstruções genômicas, provavelmente já existia no último ancestral comum das Árqueas (LACA – *Last Arquea Common Ancestor*) (Makarova et al., 2007, 2011). Curiosamente, o sistema CRISPR-Cas Tipo III é raro entre as Bactérias, encontrando-se em maior quantidade nas espécies termofílicas, o que corrobora essa teoria, podendo as Bactérias termofílicas terem recebido via transmissão horizontal os sistemas CRISPR e, posteriormente, os propagado para as demais Bactérias.

Contudo, essa teoria ainda é debatida e questionada por alguns autores. Shah et al. (2011) argumentam que são extremamente raros os casos em que transferências entre domínios sejam eficientes devido a uma série de fatores divergentes, tais como diferenças estruturais e mecanismos biológicos distintos. Assim, Shah et al. (2011) propõem que possivelmente o sistema CRISPR surgiu anteriormente à separação dos domínios, mas não descarta a possibilidade de que houve pequenas intertransferências entre Árqueas e Bactérias.

Em linhas gerais, a descoberta do sistema CRISPR-Cas é relativamente recente, de modo que a definitiva compreensão dos eventos evolutivos e moleculares que permeiam sua origem e manutenção demanda o investimento em novas pesquisas voltadas à análise detalhada da chamada “corrida armamentista evolutiva”, que atuou e ainda atua modulando a coevolução dos organismos procariotos e seus parasitas.

5. CRISPR em número de publicações

Mais de duas décadas se passaram entre a descoberta inicial do loco CRISPR (Ishino et al., 1987) e o desenvolvimento da técnica de edição genética (2012-2013) (Jinek et al., 2012; Cong et al., 2013). Poderíamos mencionar quatro fases históricas.

O período inicial de identificação e caracterização do loco CRISPR envolveu aproximadamente 20 anos (1987-2006). Nesses anos, pouca atenção foi dada ao assunto, sendo CRISPR tratado como um tópico de interesse específico da bacteriologia. De acordo com o PUBMED (<http://www.ncbi.nlm.nih.gov/pubmed>), foram publicados nesta fase 32 artigos

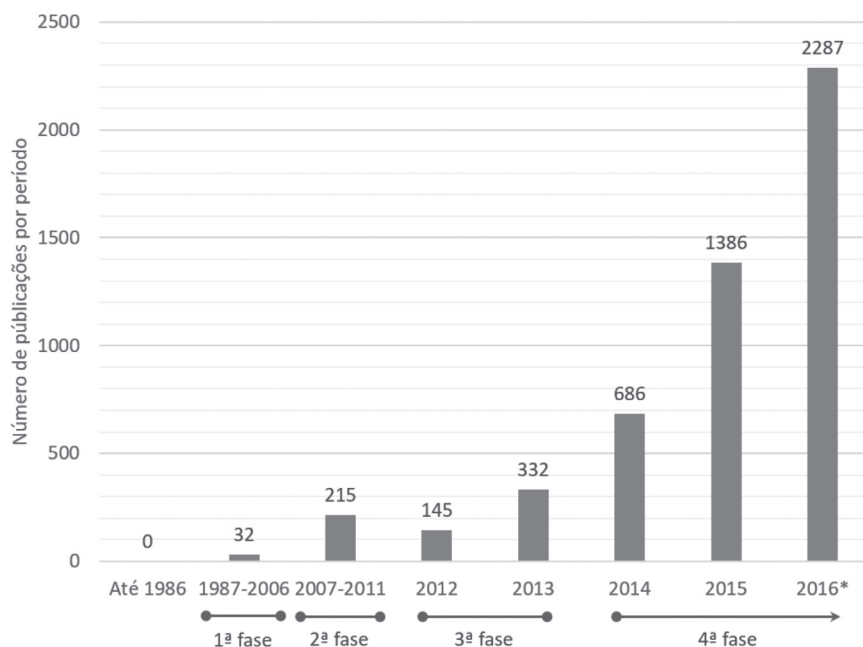


Figura 3. Crescimento no número de publicações ao longo dos anos. *Projeção para 2016. Imagem: Garcia, BH e Pereira, TC (2016).

envolvendo os termos CRISPR, *Cas9*, *Cas5*, *Csn1* ou *Csx12* (os últimos três são nomes antigos de *Cas9*; Hsu et al., 2014).

O período seguinte (2007-2011) foi marcado pelas primeiras aparições do tema em revistas de elevadíssimo impacto (*Nature*, *Science* e *Cell*), no qual a comunidade científica começava a atentar para a área e a compreender mais detalhadamente as bases moleculares do processo de resposta antiviral adquirida. Nesses anos foram publicados 215 artigos.

A fase seguinte (2012-2013) marca o nascimento da técnica de CRISPR. Há ainda algum debate sobre qual publicação, precisamente, define seu surgimento: se foi o artigo que lançou as bases e claramente evidenciou seu potencial (Jinek et al., 2012), ou o primeiro a demonstrá-la com amplo sucesso e versatilidade (Cong et al., 2013). De qualquer forma, o número de trabalhos neste biênio foi de 477.

Vivemos atualmente no quarto momento (2014-presente), caracterizado pela: (i) rápida difusão da *expertise* técnica dentro da comunidade científica, (ii) diversificação das aplicações da ferramenta, (iii) vertiginoso aumento no número de publicações e (iv) surgimento de empresas de biotecnologia baseadas em CRISPR. Mantendo-se o ritmo atual de publicações, o número de artigos apenas entre 2014-2016 deverá ser de ~4.300. Uma visão geral desses números pode ser contemplada na figura 3.

6. Termos e definições gerais

Uma série de termos e expressões são aplicados à técnica de CRISPR. Abordaremos aqui alguns dos mais elementares, para que o leitor possa ter uma melhor compreensão.

Inicialmente, é essencial separarmos o **processo biológico** da **técnica**. CRISPR é um loco bacteriano codificador de pequenos RNAs-guias de origem viral. A este loco estão associadas proteínas denominadas conjuntamente de *Cas* (*CRISPR associated*). De maneira integrada, CRISPR-Cas estão envolvidos no processo de imunidade antiviral adquirida.

Baseados nesse processo biológico natural, pesquisadores desenvolveram a técnica de **CRISPR-Cas9**, que se utiliza de um RNA-guia (referindo-se assim a CRISPR) e apenas uma das proteínas Cas (a endonuclease Cas9). Para simplificação, muitos pesquisadores referem-se à técnica simplesmente como **CRISPR**. Contudo, note que recentemente sistemas alternativos têm se utilizado de outras endonucleases que não a Cas9, como a *Cpf1* (CRISPR-Cpf1; Yamano et al., 2016).

O elemento central da técnica é uma endonuclease direcionada até o seu alvo específico por meio de um pequeno RNA-guia. Devido a isto, Cas9 é classificada como uma **RGEN** (*RNA-guided endonuclease*) (Kanchiswamy, 2016). Endonucleases são enzimas que clivagem (cortam) o ácido nucleico na parte interna da molécula, ao contrário das exonucleases, que cortam a partir das extremidades livres 5' ou 3'.

Frequentemente CRISPR é referida como uma técnica de **edição genética**, isto é, capaz de *alterar a sequência nucleotídica*. Esta alteração pode ser de diversos tipos: (i) deleção, inserção ou substituição de um ou poucos nucleotídeos, (ii) deleção de elementos genéticos, ou (iii) integração de elementos genéticos (*e.g.*, transgenes), dentre outros.

Contudo, é importante ressaltar que CRISPR pode ser usada para outros fins que não de edição genética, por exemplo: (i) marcação de DNA, (ii) regulação da expressão gênica, (iii) clivagem de RNA, (iv) mapeamento de genes ou (v) rastreamento de RNA. Nesse sentido,

CRISPR é mais do que uma técnica de edição genética. A esse conjunto versátil de técnicas derivadas os pesquisadores denominam **CRISPR toolbox** (caixa de ferramentas da CRISPR) (Lowder et al., 2015; Vora et al., 2016; Chi et al., 2016; Wright et al., 2016).

Por fim, como visto anteriormente, é curioso notar que CRISPR e RNAi compartilham mecanismos moleculares comuns: os dois processos apresentam endonucleases guiadas por pequenos RNAs, resultando na clivagem do alvo. Devido a essas e outras semelhanças, alguns grupos se referiram a CRISPR como um **RNAi-like process** natural de procariotos (Hale et al., 2008; Chakraborty et al., 2009; Terns e Terns, 2013; Kumar et al., 2015).

7. Conclusões

Ninguém imaginava que a modesta descoberta do loco CRISPR poderia desencadear uma revolução na genética molecular décadas mais tarde. Nos últimos anos, pesquisas mostraram que Cas9 é a principal nuclease envolvida no processo de clivagem de DNA invasor, resultando na utilização do sistema CRISPR como uma ferramenta fantástica de engenharia genética. Isto tornou possível editar com sucesso o genoma de células de mamíferos, assim como de inúmeros outros sistemas modelos (Hatoum-Aslan A e Marraffini LA, 2014; Jiang e Marraffini, 2015).

Referências

- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007 Mar;315(5819):1709-12.
- Barrangou R, Horvath P. CRISPR: new horizons in phage resistance and strain identification. *Annu Rev Food Sci Technol*. 2012;3:143-62.
- Barrangou R, Marraffini LA. CRISPR-Cas systems: prokaryotes upgrade to adaptative immunity. *Mol Cell*. 2014;54:234-44.
- Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet*. 2011;45:273-97.
- Bolotin A, Quinquis B, Sorokin A, Dusko Ehrlich S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*. 2005;151:2551-61.
- Chakraborty S, Waise TM, Hassan F, Kabir Y, Smith MA, Arif M. Assessment of the evolutionary origin and possibility of CRISPR-Cas (CASS) mediated RNA interference pathway in *Vibrio cholerae* O395. In *Silico Biol*. 2009;9(4):245-54.
- Chi S, Weiss A, Wang H. A CRISPR-Based Toolbox for Studying T Cell Signal Transduction. *Biomed Res Int*. 2016;2016:5052369. <http://dx.doi.org/10.1155/2016/5052369>.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013 Feb;339(6121):819-23.
- Cooper EL, Overstreet N. Diversity, evolution, and therapeutic applications of small RNAs in prokaryotic and eukaryotic immune systems. *Phys Life Rev*. 2014;11(1):113-34.
- Cooper EL. Evolution of immune systems from self/not self to danger to artificial immune systems (AIS). *Phys Life Rev*. 2010;7:55-78.
- Haerter JO, Trusina A, Sneppen K. Targeted Bacterial Immunity Buffers Phage Diversity. *J Virol*. 2011;85(20):10554-60.
- Hale C, Kleppe K, Terns RM, Terns MP. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA*. 2008 Dec;14(12):2572-9.
- Hatoum-Aslan A, Marraffini LA. Impact of CRISPR immunity on the emergence and virulence of bacterial pathogens. *Curr Opin Microbiol*. 2014;0:82-90. <http://dx.doi.org/10.1016/j.mib.2013.12.001>.
- Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*. 2010;327:167-70.
- Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014 Jun;157(6):1262-78.
- Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol*. 1987 Dec;169(12):5429-33.
- Jansen R, Embden JD, Gastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*. 2002 Mar;43(6):1565-75.

- Jansen R, Van Embden JDA, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*. 2002;43:1565-75.
- Jiang W, Marraffini LA. CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. *Annu Rev Microbiol*. 2015;69:209-28.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier EA. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012 Aug;337(6096):816-21.
- Kanchiswamy CN. DNA-free genome editing methods for targeted crop improvement. *Plant Cell Rep*. 2016 Jul;35(7):1469-74.
- Kumar MS, Plotkin JB, Hannonhalli S. Regulated CRISPR Modules Exploit a Dual Defense Strategy of Restriction and Abortive Infection in a Model of Prokaryote-Phage Coevolution. *PLoS Comput Biol*. 2015 Nov;11(11):e1004603.
- Lowder LG, Zhang D, Baltes NJ, Paul JW 3rd, Tang X, Zheng X, et al. A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol*. 2015 Oct;169(2):971-85.
- Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol Direct*. 2011;6:38.
- Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*. 2011 Jun;9(6):467-77.
- Makarova KS, Sorokin AV, Novichkov PS, Wolf YI, Koonin EV. Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol Direct*. 2007;2:33.
- Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013;31:833-8.
- Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*. 2005 Feb;60(2):174-82.
- Mojica FJ, Juez G, Rodríguez-Valera F. Transcription at different salinities of *Haloflex mediterranei* sequences adjacent to partially modified PstI sites. *Mol Microbiol*. 1993 Aug;9(3):613-21.
- Mojica FJM, Döñez-Villasenar C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol*. 2000;36:244-6.
- Sampson TR, Weiss DS. CRISPR-Cas systems: new players in gene regulation and bacterial physiology. *Front Cell Infect Microbiol*. 2014;4:37. <http://dx.doi.org/10.3389/fcimb.2014.00037>.
- Shabalina S, Koonin E. Origins and evolution of eukaryotic RNA interference. *Trends Ecol Evol*. 2008;23:578-87.
- Shah SA, Garrett RA. CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems. *Res Microbiol*. 2011;162(1):27-38.
- Stern A. CRISPR and RNA interference: similarities across immune systems: Comment on "Diversity, evolution, and therapeutic applications of small RNAs in prokaryotic and eukaryotic immune systems" by Edwin L. Cooper and Nicola Overstreet. *Phys Life Rev*. 2014;11(1):135-6.
- Terns RM, Terns MP. The RNA- and DNA-targeting CRISPR-Cas immune systems of *Pyrococcus furiosus*. *Biochem Soc Trans*. 2013 Dec;41(6):1416-21.
- Vora S, Tuttle M, Cheng J, Church G. Next stop for the CRISPR revolution: RNA guided epigenetic regulators. *FEBS J*. 2016 Jun. <http://dx.doi.org/10.1111/febs.13768>.
- Weinberger AD, Wolf YI, Lobkovsky AE, Gilmore MS, Koonin EV. Viral diversity threshold for adaptive immunity in prokaryotes. *MBio*. 2012;3(6):e00456-12.
- Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Microbiol*. 2014;12:317-26.
- Wright AV, Nuñez JK, Doudna JA. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell*. 2016 Jan;164(1-2):29-44.
- Yamano T, Nishimasu H, Zetsche B, Hirano H, Slaymaker IM, Li Y, et al. Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell*. 2016 May;165(4):949-62.

Visão geral do mecanismo básico de ação

Capítulo

2

Gabriel Viliod Vieira¹, Nerry Tatiana Cecílio¹, Letícia Magalhães Arruda¹
e Katiuchia Uzzun Sales¹

¹Dept.º de Biologia Celular e Molecular, FMRP, USP, SP - Brasil

Estruturação do capítulo

1. Visão geral sobre moléculas envolvidas e mecanismos básicos de ação
 2. Moléculas envolvidas na técnica de CRISPR
 - 2.1 A nuclease Cas9
 - 2.2 RNA-guia (*Single guide RNA*)
 - 2.3 DNA-alvo
 3. Mecanismo de ação de Cas9/sgRNA
 - 3.1 Clivagem de corte duplo
 - 3.2 Clivagem de corte simples
 4. Reparo do DNA: NHEJ e HDR
 - 4.1 Junção de Pontas Não Homólogas (NHEJ)
 - 4.2 Reparo Direcionado por Homologia (HDR)
 5. Conclusão
- Apêndices

1. Visão geral sobre moléculas envolvidas e mecanismos básicos de ação

Na natureza, CRISPR-Cas é um sistema procariótico de imunidade baseado na captura e inserção de pequenos pedaços de DNA advindos da invasão por vírus ou plasmídeos, que são incorporados ao genoma da bactéria e contra os quais ela então adquire resistência (Horvath e Barrangou 2010).

Com base nesse mecanismo natural foi desenvolvida a *técnica* de CRISPR. Para o entendimento do papel das moléculas envolvidas nessa tecnologia, deve-se levar em consideração a plasticidade dos ácidos nucleicos em se dobrarem, formando estruturas secundárias, e interagirem entre si, nas seguintes formas: (i) DNA:DNA, (ii) RNA:RNA e (iii) DNA:RNA. Adicionalmente, o DNA e (ou) RNA podem se associar com outras classes de moléculas, como as proteínas.

Basicamente, a técnica de CRISPR envolve três moléculas: uma **nuclease** (geralmente a Cas9 tipo selvagem de *Streptococcus pyogenes*), um **RNA-guia** (conhecido como *single guide* RNA) e o **alvo** (frequentemente o DNA).

2. Moléculas envolvidas na técnica de CRISPR

2.1 A nuclease Cas9

Cas9 é uma proteína bacteriana pertencente ao sistema CRISPR do Tipo II (vide capítulo 1), capaz de clivar a dupla fita de DNA em sua parte interna (*i.e.*, uma endonuclease de corte duplo). A análise tridimensional da Cas9 de *Streptococcus pyogenes* (SpCas9), que é a enzima mais comumente utilizada na edição do DNA de células eucarióticas, revelou que ela possui estrutura bilobulada: um lóbulo de reconhecimento (**REC**) e um lóbulo com atividade nucleásica (**NUC**). O lóbulo NUC consiste nos *domínios catalíticos* **RuvC** e **HNH**, além do *domínio de interação com PAM* (**PI**) (Jinek M. et al., 2012).

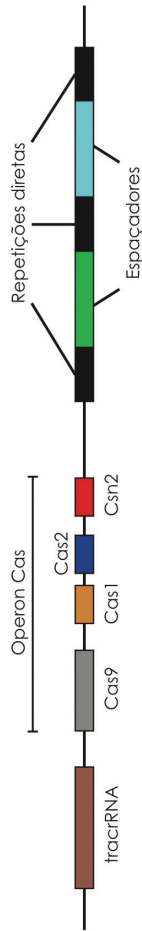
2.2 RNA-guia (*Single guide* RNA)

Para que Cas9 possa executar sua função, ela precisa ser *ativada* e *direcionada* ao seu alvo. No processo de imunidade natural em bactérias, estas etapas são mediadas por duas moléculas de RNA que atuam cooperativamente: o crRNA (*CRISPR-derived RNA*) e tracrRNA (*trans-activating RNA*) (figura 1) (vide apêndice 1 para as características da imunidade bacteriana, crRNA e tracrRNA).

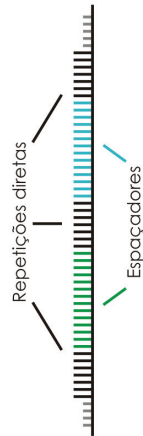
Objetivando tornar este processo o mais simples possível para a aplicação em laboratório, pesquisadores desenvolveram o ***single guide* RNA (sgRNA ou gRNA)**: uma molécula química resultante da “fusão” do crRNA e tracrRNA, gerada para acumular as duas funções, que são altamente dependentes de suas estruturas (figuras 1D e 1E). Dessa forma, a técnica de CRISPR demanda apenas dois elementos exógenos: a Cas9 e sgRNA, facilitando assim o procedimento experimental.

O sgRNA é composto por um dobramento em grampo formado pela *sequência de reconhecimento do alvo* também conhecida como sequência-**guia** (~20 nt na extremidade 5', específico para cada alvo) mais uma **sequência universal** (~80 nt, o *scaffold*, extremidade 3')

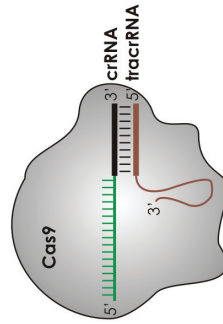
A) Loco CRISPR



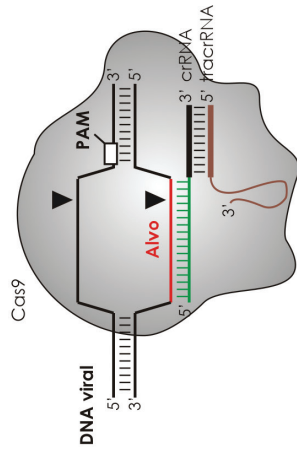
B) Formação do pré-crRNA



C) Maturação do pré-crRNA pela RNase III e formação do complexo com Cas9



D) Clivagem do DNA por Cas9



E) Clivagem do DNA por Cas9/gRNA

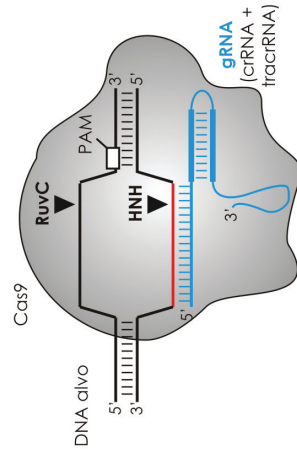


Figura 1. CRISPR-Cas9 no sistema imune bacteriano e na edição gênica. A. Loco genômico bacteriano CRISPR-Cas9, evidenciando os genes que o compõem (tracrRNA, Cas9, Cas1, Cas2, Casn2, Repetições diretas e Espaçadores). B. Transcrição do pré-crRNA. C. Pré-crRNA sofre processamento e maturação pela RNase III, além de outras nucleases, permitindo a formação de um complexo contendo: a nuclease Cas9, o tracrRNA e o crRNA. D. O complexo formado é capaz de efetuar a clivagem do DNA exógeno de origem viral, dependente da interação com a sequência PAM. E. Complexo formado pela Cas9 e sgRNA, utilizado para a edição gênica em células de mamíferos, dependente da interação com a sequência PAM. Imagem: Pereira, TC (2016).

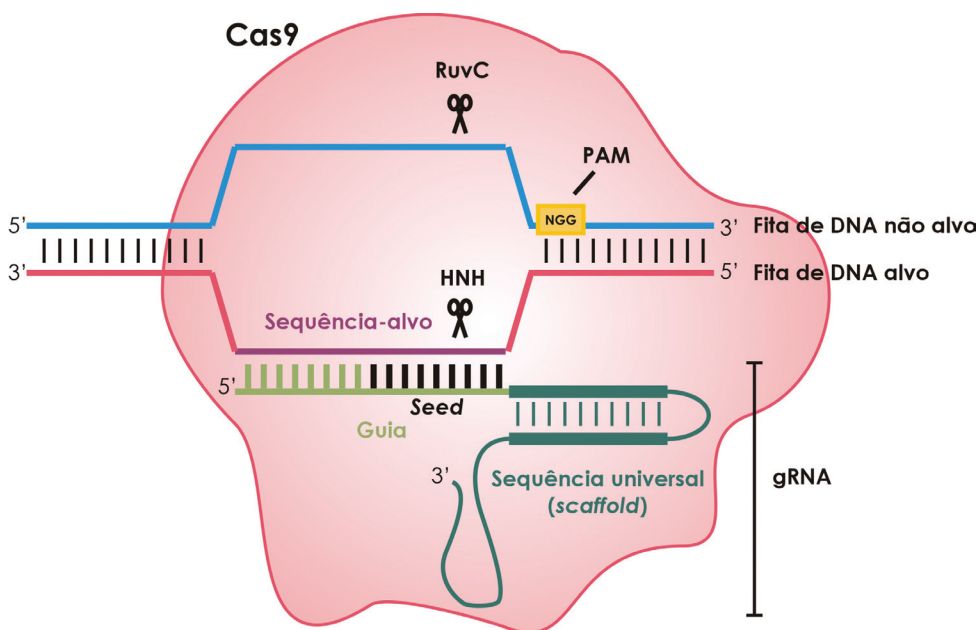


Figura 2. Representação esquemática do complexo Cas9/sgrRNA/DNA-alvo. O gRNA está representado em verde, ressaltando a sequência universal (*scaffold* – verde-escuro), guia (verde-claro) e a *seed* (em preto, faz parte da guia). O DNA-alvo, complementar ao guia, está representado em rosa e a fita não complementar representada em azul, com a sequência PAM destacada em amarelo. A sequência-alvo está indicada em roxo. A enzima Cas9 está representada ao fundo em rosa-claro. Os domínios catalíticos RuvC e HNH responsáveis pela clivagem da fita não complementar e complementar, respectivamente, estão indicados pelas tesouras. Imagem: Pereira, TC (2016).

e conserva as interações de pareamento das bases na dupla fita (figura 2) (Jinek et al., 2012). O sgRNA teve sua aplicabilidade inicialmente testada em procariotos (Doudna e Charpentier 2014; Jinek et al., 2012) e tem sido extensivamente empregado na edição de genomas de células de mamíferos (Cong et al., 2013).

2.3 DNA-alvo

A região no DNA que será clivada pela nuclease Cas9 possui dois elementos: (i) o **alvo** propriamente dito e (ii) uma sequência **PAM** (*Protospacer Adjacent Motif*) (figuras 1E e 2) (Deveau, H. et al., 2008); (Garneau et al., 2010).

Como mencionado anteriormente, o sgRNA possui em sua extremidade 5' uma sequência-*guia* (~20 nt), a qual irá direcionar Cas9 a encontrar uma sequência-*alvo* complementar de ~20 nt no DNA. Essa característica tornaria o evento teoricamente específico, uma vez que um alvo de 20 nt ocorre apenas uma vez a cada um trilhão de nt (4^{20}). Contudo experimentalmente, eventos inespecíficos (*off-targets*) são reportados. Alguns dos motivos desses erros serão discutidos posteriormente.

O complexo Cas9/sgrRNA irá interagir com o *alvo* apenas se houver um motivo PAM adjacente na outra fita de DNA (figura 2). PAMs são sequências curtas, geralmente de 2-5 nt (*e.g.*, 5' NGG3' e 5' NNGRRT3'), essenciais para a ancoragem da nuclease ao sítio de clivagem.

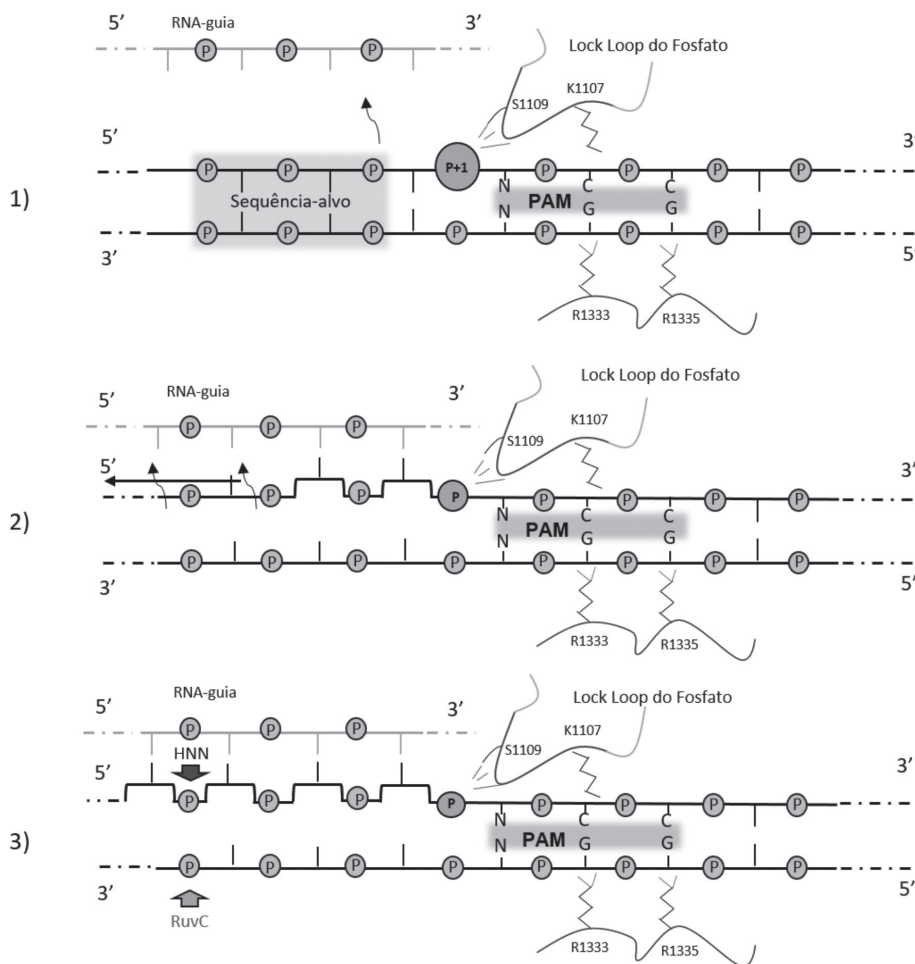


Figura 3. Representação esquemática do reconhecimento da sequência PAM pela Cas9 e organização estrutural do complexo de reconhecimento. 1. **Ligação a PAM:** A sequência PAM é essencial para o reconhecimento pela Cas9. Os resíduos de aminoácidos Arginina 1333 e 1335 de Cas9, contidos no domínio de Interação a PAM (PI), interagem com as duas bases guanina (NGG) do sulco maior do DNA. Já no sulco menor é o resíduo Lisina 1107 que interage com a PAM. 2. **Lock Loop do fosfato:** Esse laço consegue posicionar o DNA de forma a permitir que o resíduo de Serina 1106 da PAM e dois nitrogênios, presentes no laço, formem ligações de hidrogênio na posição +1 da sequência PAM. Esse processo permite a estabilização do DNA-alvo, direcionando as primeiras bases nitrogenadas em direção ao RNA-guia. 3. **RNA-guia:** O DNA-alvo se pareia então com o RNA-guia. Caso não houvesse a ligação inicial pela PAM e a estabilização do fosfato +1, não seria possível observar a interação do RNA-guia com o DNA-alvo. Depois da correta ligação à sequência PAM, interação do RNA-guia com o DNA-alvo e estabilização de todo o complexo, ocorre a clivagem da dupla fita do DNA pelos domínios HNH e RuvC da Cas 9 (indicado pelas setas). Imagem: Vieira, GV; Cecílio, NT; Arruda, LM e Sales, KU (2016).

3. Mecanismo de ação de Cas9/sgRNA

3.1 Clivagem de corte duplo

Pesquisas de cristalografia da SpCas9 evidenciaram algumas das principais etapas do processo de clivagem. Inicialmente, a interação de Cas9 com o sgRNA resulta em uma alteração na estrutura da proteína, com a formação do *domínio de interação com a PAM*

(PI) e conseguinte associação do complexo Cas9/sgRNA ao DNA-alvo (por meio de PAM). Nesta configuração, PI atua simultaneamente nas duas fitas do DNA-alvo, mecanicamente separando-as (figura 3).

Isto ocorre da seguinte maneira: resíduos conservados de arginina no PI ligam-se ao dinucleotídeo “GG” do motivo PAM (5’ NGG3’), espacialmente limitando a movimentação da cadeia não alvo. Concomitantemente, outra região do PI interage com um grupamento fosfodiéster na cadeia-alvo, “afastando-a” da cadeia não alvo. Esta ação resulta na separação *local* das cadeias do DNA-alvo (abertura das fitas) imediatamente a montante da sequência PAM.

Esta abertura local permite, por conseguinte, iniciar a interação da sequência-*guia* (do sgRNA) com o *alvo* em si, que ocorre logo a montante de PAM (Anders et al., 2015). Em especial, esta etapa inicial de interação envolve um domínio fundamental da sequência-*guia* chamado *seed*, de ~8 nt de extensão (figura 2) (Semenova et al., 2011). A partir da *seed*, o pareamento entre sgRNA e alvo é iniciado (ou *semeado*). Mal pareamentos entre *seed* e um possível alvo não são tolerados, impedindo assim a estabilização de Cas9/sgRNA com esta molécula. De outro modo, havendo perfeito pareamento entre *seed* e o alvo, o processo de abertura (separação das duas fitas de DNA) prossegue, resultando em uma interação estável entre Cas9/sgRNA e o alvo. Ao final, a carga negativa do heterodúplex sgRNA/DNA-alvo é acomodada no sulco formado na interface entre os lóbulos REC e NUC de Cas9, que apresentam carga positiva.

Nesse momento, Cas9 está habilitada para clivar o alvo nas duas fitas de DNA, por meio de seus domínios catalíticos HNH (que cliva a fita-alvo, complementar ao guia) e RuvC (que cliva a outra fita). Esta clivagem ocorre dentro do *alvo*, cerca de 3-4 pb a montante de PAM, gerando extremidades abruptas (*i.e.*, *blunt-end*) (Jinek et al., 2012).

3.2 Clivagem de corte simples

É importante mencionar que, a partir da **Cas9 selvagem** (*i.e.*, não modificada), uma *variante* foi recentemente desenvolvida, denominada Cas9 ***nickase*** (Cas9n). Esta versão é capaz de realizar corte em apenas uma das duas cadeias do DNA-alvo, devido à inativação de um de seus domínios catalíticos (RuvC ou HNH) (Gasiunas et al., 2012; Jinek et al., 2012; Sapranasauskas et al., 2011). As características, particularidades e vantagens do uso de *nickases* serão discutidas no capítulo 6.

4. Reparo do DNA: NHEJ e HDR

As clivagens realizadas por Cas9, gerando extremidades abruptas (corte em ambas as fitas) ou *nicks* (corte em apenas uma das fitas), devem ser reparadas pela célula. Neste cenário, dois mecanismos podem entrar em ação: (i) a via de ***junção de pontas não homólogas*** (NHEJ em inglês, *non-homologous end joining*) ou (ii) a via de ***reparo direcionado por homologia*** (HDR em inglês, *homology directed repair*).

4.1 Junção de Pontas Não Homólogas (NHEJ)

Como visto anteriormente, a atividade da Cas9 selvagem gera extremidades abruptas, isto é, pontas de DNA que geralmente não possuem homologia (ou complementaridade)

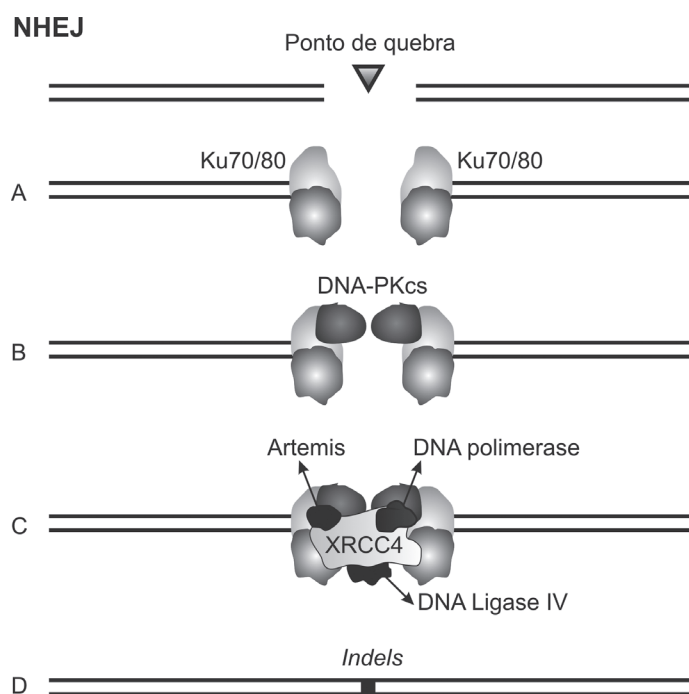


Figura 4. Reparo por Junção de Pontas Não Homólogas (NHEJ). A. Ligação das proteínas Ku 70 e 80 nas extremidades da quebra de dupla fita do DNA, formando um heterodímero. B. Proteínas Ku servem como ancoragem para a ligação de DNA-PKcs. C. Outras proteínas passam a fazer parte do complexo, como Artemis, XRCC4, DNA polimerase e DNA Ligase IV, efetuando o processamento das extremidades. D. Reparo do DNA finalizado pela DNA Ligase IV. Ao final do processo a dupla fita de DNA apresenta inserção/deleção de alguns nucleotídeos (*indels*). Imagem: Pereira, TC (2016).

entre si (figura 4). Esta quebra da dupla fita de DNA (DSB em inglês, *double strand break*) representa um dano importante às células eucarióticas e procarióticas, pois pode levá-las à apoptose ou senescência. O reparo deste dano é feito pela via NHEJ, que é a mais comum em células de mamíferos e ocorre em qualquer etapa do ciclo celular (vide apêndice 2 para detalhes de NHEJ).

Em essência, este mecanismo irá reunir as duas pontas de DNA, por meio de eventos não dependentes de homologia entre as duas extremidades. Curiosamente, esta via é intrinsecamente propensa ao erro, gerando mutações do tipo *indel* (inserção ou deleção de um ou poucos nucleotídeos) no sítio de junção das pontas, ou próximo a ele (Doudna e Charpentier, 2014). Estas mutações, por sua vez, comprometerão a funcionalidade do produto final do gene-alvo. Ou seja, a ação conjunta de Cas9-NHEJ frequentemente resulta no **nocauteamento** (inativação) do gene. O nocaute gênico é de grande interesse para a genética, biologia celular, biomedicina e diversas outras áreas, apresentando diversas aplicações (vide capítulo 4).

4.2 Reparo Direcionado por Homologia (HDR)

Em alguns casos, as extremidades geradas pela clivagem apresentam homologia entre si (ou com alguma terceira molécula). Nessas situações, outra via de reparo poderá ser recrutada – a HDR, baseada em recombinação homóloga. Nas células eucarióticas,

observamos três principais formas diferentes de recombinação homóloga que podem ser utilizadas no reparo por HDR: (i) a conversão gênica, (ii) o reparo induzido por quebra e (iii) o anelamento de uma única fita de DNA.

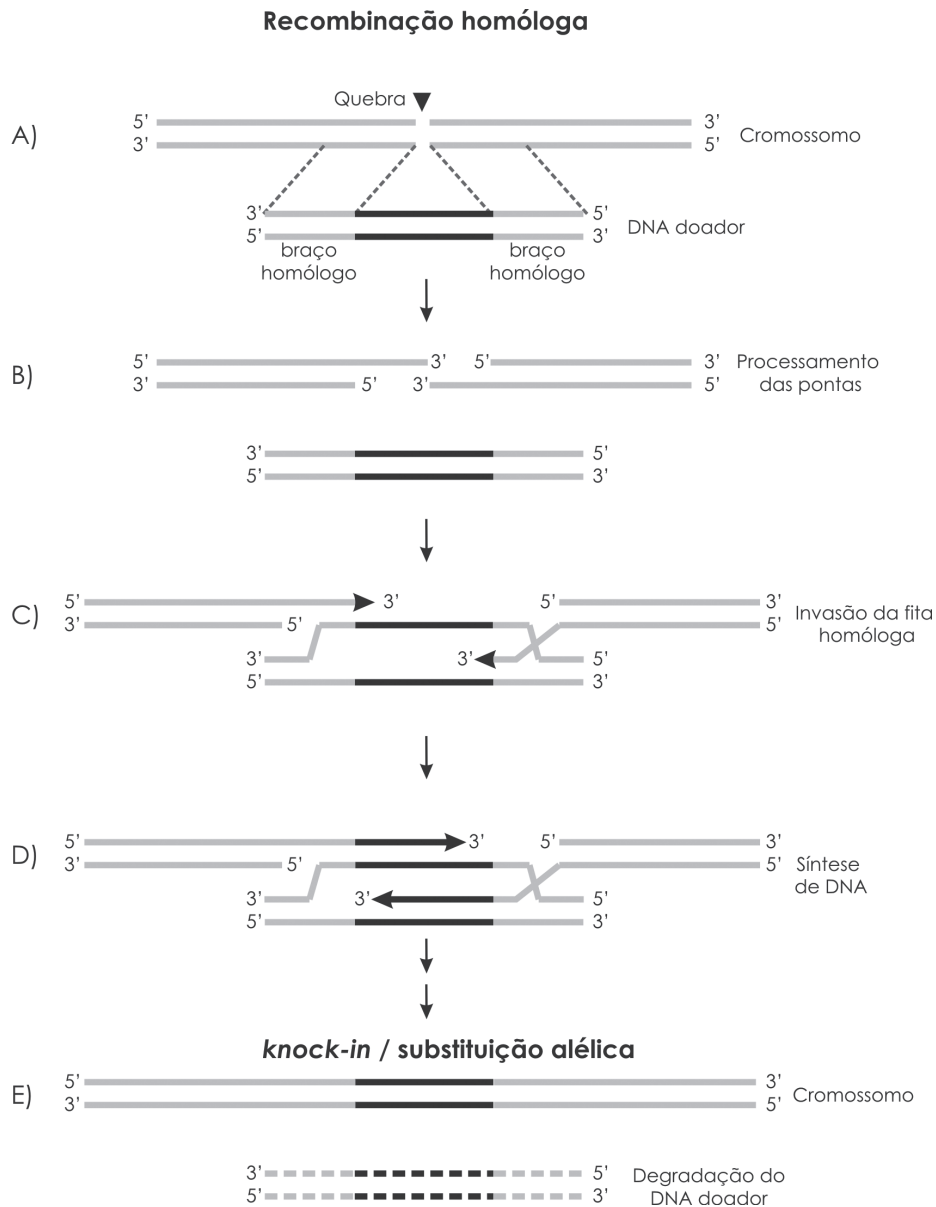


Figura 5. Recombinação homóloga resultando em *knock-in*/substituição alélica. A quebra de dupla fita do DNA pode ser reparada via HDR na presença de um DNA homólogo. A) O DNA doador possui braços de homologia com a região adjacente ao sítio de clivagem no cromossomo. B) Estas pontas do cromossomo sofrem processamento, deixando extremidades 3' livres. C) Posteriormente ocorre a invasão de uma fita do DNA na região homóloga do DNA doador e (D) síntese de DNA. E) O resultado deste processo é a integração de uma sequência exógena no genoma (*knock-in*) ou substituição alélica. Imagem: Pereira, TC (2016).

Contudo, frequentemente a clivagem do DNA-alvo por Cas9 não gera extremidades com algum grau de homologia. Mesmo assim, a HDR pode ser *desencadeada* caso se introduza no experimento uma molécula de **DNA doador** (*i.e.*, a *terceira* molécula mencionada no parágrafo anterior), a qual deve possuir homologia com *ambas as extremidades* resultantes do corte (figura 5).

O desencadeamento da HDR é essencial quando se objetiva gerar células/organismos “**knock-in**” ou realizar **substituição alélica** (figura 5). *Knock-in* geralmente se refere à introdução de uma nova sequência de DNA no genoma, como um transgene por exemplo (Tom Strachan e Andrew Read, 2013). A substituição alélica refere-se à troca de uma sequência nucleotídica por outra muito semelhante, porém distinta (*i.e.*, de um alelo por outro) (Zhang et al., 2014). Em ambos os casos, o procedimento experimental de CRISPR é modificado, agora três elementos exógenos são introduzidos na célula: a nuclease Cas9, o sgRNA e o DNA doador (o desencadeador da HDR).

Nesse contexto, Cas9 irá realizar o corte em um ponto específico no genoma, ao qual o DNA doador apresenta homologia. Assim, logo após a clivagem o HDR irá promover o reparo deste dano usando como molde o DNA doador. Se o DNA doador apresentar uma variação nucleotídica em relação ao DNA-alvo (como um *Single Nucleotide Polymorphism* – SNP), o evento resultará em uma substituição alélica. Contudo, se o DNA doador for um transgene flanqueado por sequências homólogas ao sítio de clivagem, o reparo resultará na integração sítio-específica do transgene. É importante ressaltar que, atualmente, a HDR tem sido predominantemente realizada utilizando-se *nickases*.

5. Conclusão

Originalmente descoberto como um sistema de defesa bacteriano, CRISPR tornou-se a mais importante ferramenta em engenharia genética desta década. Baseada em um conceito relativamente simples, mas não menos elegante, a tecnologia CRISPR utiliza-se de uma nuclease (Cas9) guiada por um RNA (sgRNA) ao DNA-alvo (contendo uma sequência PAM), resultando na clivagem da dupla fita de DNA em um loco específico. Esta quebra do DNA pode, então, ser reparada por: (i) junção de pontas não homólogas (NHEJ), que acontece predominantemente a partir de quebras geradas pela Cas9 selvagem e na ausência de DNA doador ou; (ii) reparo direcionado por homologia (HDR), que acontece predominantemente a partir de rupturas geradas por Cas9 *nickases* e na presença de DNA doador. O reparo por junção de pontas não homólogas promove, na maioria das vezes, mutações *indel* (inserções ou deleções). Estas, por sua vez, culminam em modificação da fase de leitura do DNA (*frameshift*) e em um códon de parada prematuro, resultando na geração de organismos ou células nocauteadas. Por outro lado, o reparo direcionado por homologia culmina principalmente em inserções de transgenes ou substituições alélicas no genoma, de acordo com a necessidade do estudo, o que chamamos de *knock-ins*.

Apêndice 1

Os pequenos pedaços de DNA (de origem viral, plasmidial ou de transposons) incorporados ao DNA genômico da bactéria denominam-se espaçadores (dentro dos DNAs

clusterizados), e são alojados entre sequências repetidas, dentro dos operons associados aos genes Cas (*CRISPR-associated*), dispostos em *tandem* (Barrangou et al., 2007) (figura 1A). O conjunto de espaçadores é transcrito em um único RNA (pré-crRNA) (figura 1B), que é processado a crRNA, em associação com: (i) o *trans-activating* RNA (tracrRNA), (ii) a Cas9 e (iii) a ribonuclease III (RNase III) (figura 1C). Dessa forma, as sequências dos crRNAs maduros corresponderão à transcrição dos pequenos pedaços de DNAs virais incorporados ao genoma bacteriano. Posteriormente, o crRNAs maduros, juntamente com a proteína Cas9 e o *trans-activating* RNA (tracrRNA), são capazes de silenciar DNAs exógenos de maneira muito específica (figura 1D) (Jinek et al., 2012).

O **crRNA** é o RNA que reconhece o DNA-alvo. No sistema bacteriano natural, é sintetizado na forma de pré-crRNA durante a transcrição do operon dos genes *cas* (Barrangou et al., 2007). A edição do pré-crRNA em crRNA é realizada pela enzima RNase III, após a sua associação com o tracrRNA (Deltcheva et al., 2011). A RNase III é um fator geral de processamento de RNA bacteriano, recrutado na presença de dúplice de RNA e no sistema bacteriano CRISPR Tipo II, é estritamente necessária na maturação do crRNA. Para que seja ativo, o crRNA maduro deve se associar à fita complementar do DNA-alvo na região adjacente à PAM, de forma que, juntamente com o tracrRNA e a Cas9, forme o complexo efetor do sistema CRISPR-Cas (Ran, F.A. et al., 2013).

O **tracrRNA** é um pequeno RNA não codificante complementar às sequências repetidas do transcrito precursor do crRNA (pré-crRNA) e auxilia na sua maturação por intermédio da formação de um híbrido entre o tracrRNA e o pré-crRNA (Deltcheva et al., 2011). Ainda, o pareamento tracrRNA/crRNA é necessário para que haja o correto reconhecimento e direcionamento do DNA-alvo na interação do DNA com a Cas9. Em termos estruturais, a porção inicial 5' do tracrRNA se associa à porção terminal 3' do crRNA, de forma que as outras extremidades ficam livres para outras associações: a extremidade 5' do crRNA atua no reconhecimento do DNA-alvo e a porção 3' do tracrRNA pode interagir com a Cas9. Em resumo, a formação da dupla fita de RNA, na interação tracrRNA/crRNA, é o cerne necessário para que o crRNA reconheça o DNA-alvo e para a atividade catalítica da Cas9.

Apêndice 2

O NHEJ pode ser didaticamente dividido em quatro etapas (Anthony et al., 2013):

- a. O processamento inicial da quebra de dupla fita, comum aos processos de HDR e NHEJ, pelo complexo MRN (formado pelas proteínas MRE11, RAD50 e Nibrina), continua através do reconhecimento das extremidades danificadas do DNA por duas proteínas essenciais: Ku70 e Ku80 (figura 4A), que funcionam como ancoragem para outras proteínas importantes durante a NHEJ.
- b. A formação do heterodímero composto pelas proteínas Ku 70 e 80 serve como suporte para que outras proteínas, como a DNA-PKcs, XRCC4 e APLF, sejam recrutadas para a área danificada do DNA proporcionando estabilidade e, assim, impeçam o processamento não específico desta molécula (figura 4B).
- c. A molécula de DNA danificada é, então, processada, para ser posteriormente religada. O processamento de suas extremidades é feito por diversas enzimas, como Artemis, PNKP, APLF, Polimerase μ , entre (figura 4C).

- d. Extremidades do DNA são, assim, religadas pela ação da DNA Ligase IV. Diversas outras proteínas são necessárias para estabilizar o processo. Vale a citação da proteína XRCC4, que atua como um cofator essencial para a DNA Ligase IV.

Referências

- Anders C, et al. Europe PMC Funders Group Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. 2015;513(7519):569-73.
- Barrangou R, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*. 2007;315:1709-12.
- Cong L, et al. Multiplex Genome Engineering Using CRISPR / Cas Systems. *Science*. 2013;339:819-23.
- Deltcheva E, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 2011;471(7340):602-7.
- Deveau H, et al. Phage Response to CRISPR-Encoded Resistance in *Streptococcus thermophilus*. *J Bacteriol*. 2008;190(4):1390-400.
- Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346(6213):1258096-9.
- Garneau, JE. et al. The CRISPR / Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010;468(7320):67-71.
- Gasiunas, G. et al. Cas9 – crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. 2012;109(39):2579-86.
- Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*. 2010;327(5962):167-70.
- Jinek M, et al. A Programmable Dual-RNA – Guided. *Science*. 2012;337:816-22.
- Ran FA, et al. Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell*. 2013;154(6):1380-9.
- Ran FA, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-308.
- Sapranauskas R, et al. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res*. 2011;39(21):9275-82.
- Semenova E, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA*. 2011;108(25):10098-103.
- Strachan T, Read A. *Genética molecular humana*. 4. ed. Artmed; 2013.
- Zhang C, Xiao B, Jiang Y, Zhao Y, Li Z, Gao H, et al. Efficient editing of malaria parasite genome using the CRISPR/Cas9 system. *MBio*. 2014 Jul;5(4):e01414-14.

CRISPR *toolbox*: a versatilidade da técnica

Capítulo 3

Caroline Junqueira^{1,a}, Patrícia Renck Nunes^{2,a}, Gabriela Molinari Roberto^{3,b}
e María Sol Brassesco Annichini^{4,b}

¹*Centro de Pesquisas René Rachou/ FIOCRUZ, MG - Brasil*

²*Dept.º de Ciências da Vida, Escola Politécnica Federal de Lausanne - Suíça*

³*Programa de Pós-Graduação em Oncologia Clínica, FMRP USP, SP - Brasil*

⁴*Dept.º de Biologia, FFCLRP USP, SP - Brasil*

^a*Autoria das seções: 1 a 5*

^b*Autoria das seções: 6 e 7*

Estruturação do capítulo

1. Introdução
2. Edição de DNA
 - 2.1 Inativação gênica
 - 2.2 Integração de transgenes
 - 2.3 Substituição alélica
 - 2.4 Deleção de genes inteiros
3. Marcação de DNA
4. Regulação da expressão gênica
 - 4.1 Para indução da expressão gênica
 - 4.2 Para repressão da expressão gênica
5. Clivagem de RNAs
6. Mapeamento de genes
7. Rastreamento de RNAs
8. Conclusões

1. Introdução

Em essência, a versão inicial da técnica de CRISPR promovia uma clivagem sítio dirigida no DNA-alvo, que poderia ser reparado por NHEJ (culminando em mutações pontuais aleatórias na sequência-alvo) ou HDR (resultando em alterações precisas na sequência nucleotídica-alvo).

Nesta seção veremos como algumas alterações no protocolo experimental permitiram tornar a CRISPR mais versátil, expandido seu repertório de atuação além da **edição genética** (*gene editing*) tais como: **marcação de DNA** (*DNA labeling*), **regulação da expressão gênica** (indução e repressão gênicas), **clivagem de RNAs**, **mapeamento de genes** e **rastreamento de RNA** (*RNA tracking*). As aplicações de cada uma delas serão vistas no capítulo seguinte.

2. Edição de DNA

No contexto da técnica de CRISPR, a expressão “**edição genética**” se refere à capacidade de *alterar a sequência nucleotídica do gene-alvo*, em analogia simples à *edição de textos* no computador: alteramos algumas letras (nucleotídeos) ou palavras (genes inteiros) de um documento (material genético), afetando assim seu significado. Quando estas alterações são *muito extensas* (envolvendo regiões cromossômicas) ou *muito numerosas* (englobando múltiplos genes simultaneamente), referimo-nos à “**edição genômica**”.

As diversas formas de edição genética podem ser categorizadas de várias maneiras. Aqui adotaremos uma classificação que é relativamente comum e simples: (i) a inativação gênica, (ii) a integração de transgenes, (iii) a substituição alélica e (iv) a deleção de genes inteiros.

2.1 Inativação gênica

A inativação ou nocauteamento gênico (*gene knockout - KO*) é tipicamente alcançado por meio de mutações que resultam na *perda da atividade do produto final do gene* (seja ele um RNA funcional ou uma proteína) (William S. Klug et al., 2009). A consequência dessas alterações genéticas são alelos **nulos**, *i.e.*, sem função biológica (Benjamin Lewin, 2008). O nocauteamento pode ser alcançado tanto por clivagem por Cas9 seguida por NHEJ quanto pela HDR, sendo esta última mais específica, por induzir alterações precisas.

Como descrito no capítulo anterior, NHEJ é um processo propenso ao erro, promovendo *indels* no sítio de clivagem (figura 1). Esta inserção/deleção de um ou poucos nucleotídeos, quando não múltiplos de três, altera a *fase de leitura* do gene em questão. Isto, por sua vez, quase sempre resulta em um *códon de parada prematuro*, gerando assim proteínas truncadas (figura 1) (Hsu, et al., 2014; Mei et al. 2016).

Apesar de sua vasta aplicação para o nocauteamento de genes, NHEJ não permite um controle da mutação gerada. Eventualmente podem ocorrer inserções/deleções de nucleotídeos em um número múltiplo de três e *em fase* com a matriz de leitura, resultando na simples adição ou deleção de um ou poucos aminoácidos na proteína final. Esta pequena alteração pode ou não ter efeito biológico significativo.

Alternativamente, pode-se obter o nocaute gênico por HDR, que permite modificações precisas no gene-alvo na presença de um DNA doador. Neste cenário, o DNA doador é, em si, a *versão nocauteada* do gene-alvo. Esta versão nocauteada deve ser construída em laboratório, apresentar uma mutação bem definida e será integrada no genoma por recombinação (figura 2). Para tanto, faz-se necessário a cotransfecção de Cas9, gRNA e o DNA doador. Para edições

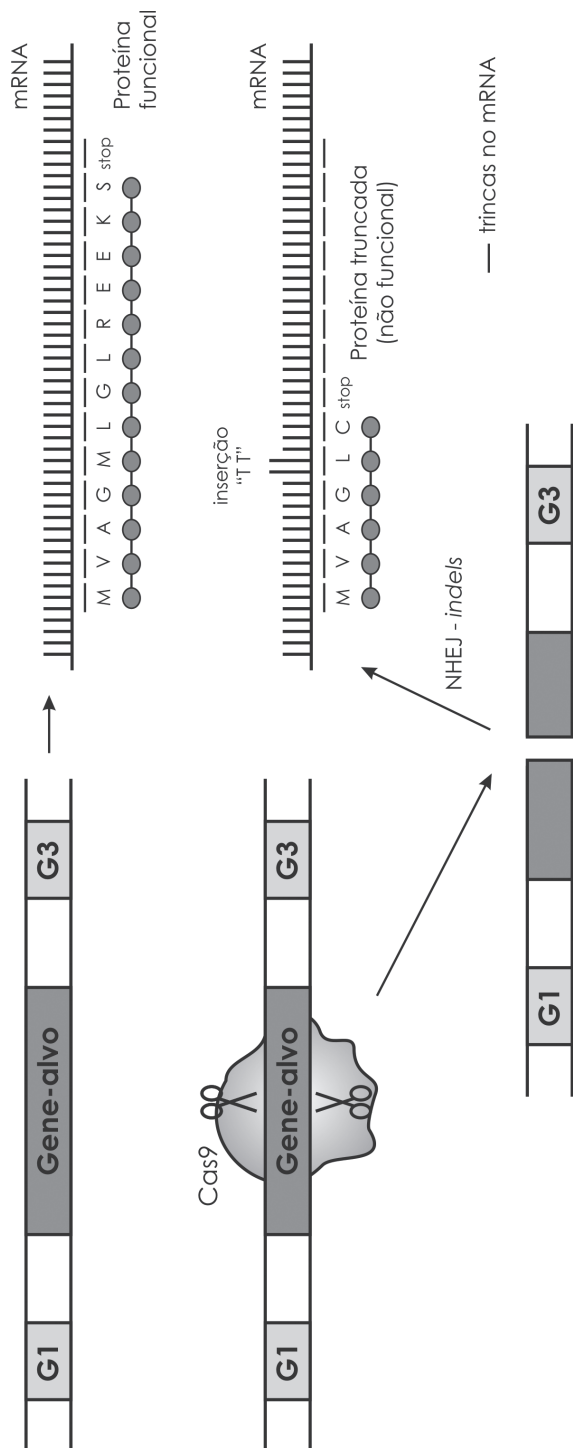


Figura 1. Nocauteamento gênico via CRISPR. A clivagem por Cas9 seguida por NHEJ gera *indels* que, quando não múltiplos de três nucleotídeos, alteram a fase de leitura resultando em códons prematuros de parada. Este polipeptídeo truncado frequentemente não é funcional. G1 e G3: “gene 1” e “gene 3”, para referência. Imagem: Pereira, TC (2016).

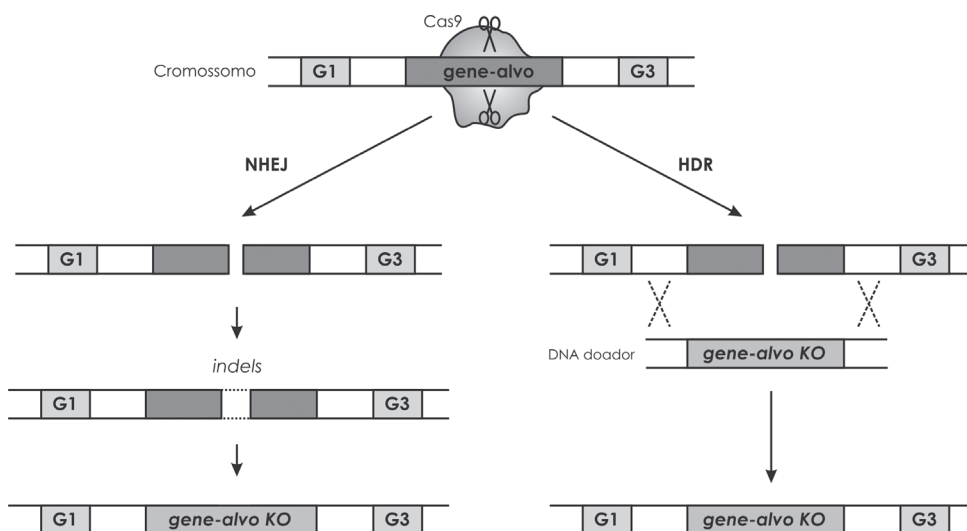


Figura 2. Mecanismos de nocauteamento via CRISPR. Após a clivagem por Cas9, NHEJ repara a dupla fita de DNA, adicionando ou deletando nucleotídeos (*indels*). Esta mutação pontual tem como consequência o aparecimento de um códon prematuro de parada, levando ao nocauteamento do gene (KO). Alternativamente, HDR repara a dupla fita de DNA por recombinação homóloga na presença de um DNA doador com extremidades homólogas ao DNA original (versão nocauteada do gene), usando-o como molde para recombinação. Imagem: Pereira, TC (2016).

curtas, de até 50 pares de bases (bp), oligonucleotídeos de fita simples podem ser utilizados como DNA doador, sendo uma opção simplificada (vide capítulo 9 para detalhes experimentais).

Estratégias mais elaboradas podem ser feitas para o nocauteamento via HDR. Por exemplo, o DNA doador pode ser um *gene marcador de resistência a antibióticos* (e.g., Neo^R) ou um *gene codificador de uma proteína fluorescente* (e.g., GFP) com extremidades homólogas ao gene-alvo. Como resultado da HDR temos uma inativação insercional, i.e., a inserção de Neo^R (ou GFP) dentro do gene-alvo resultando em seu nocauteamento. Adicionalmente, a seleção das células modificadas por esta abordagem torna-se um procedimento extremamente fácil.

Contudo, ao contrário de NHEJ, HDR é geralmente ativa apenas em células em divisão. Isto ocorre porque as células possuem mecanismos de inibição de HDR durante as fases iniciais do ciclo celular, a fim de evitar fusão de telômeros. Adicionalmente, a eficácia de HDR pode variar amplamente dependendo do tipo celular, do local de edição e do modelo de reparo. Portanto, apesar de NHEJ não ser o modelo de nocauteamento genético mais preciso, é por muitas vezes a técnica de escolha devido a dificuldades metodológicas de implementação de HDR (RAN et al., 2014).

Em síntese, a inativação gênica (KO) pela técnica de CRISPR pode ser alcançada via NHEJ, gerando *indels* que frequentemente resultam em proteínas truncadas. Alternativamente, o KO pode ser obtido via HDR, utilizando um DNA doador, levando a uma inativação de natureza predeterminada. Todas estas modificações resultam em alelos nulos.

2.2 Integração de transgenes

Apesar de existirem diversas estratégias para a integração de sequências exógenas (e.g., transgenes), geralmente elas não são sítio dirigidas, i.e., o evento ocorre ao acaso no

genoma. Isto pode ser problemático se a inserção do transgene ocorrer *dentro* de um gene endógeno importante, inadvertidamente inativando-o.

A estratégia de *knock-in* (KI) foi desenvolvida para sanar esta problemática, permitindo a integração de transgenes em regiões predeterminadas do genoma, evitando-se assim efeitos indesejados (Tom Strachan e Andrew Read, 2013). Isto é possível porque a KI é um processo baseado em recombinação homóloga (HDR), muito semelhante ao visto para KO. A única diferença é que o DNA doador, neste caso, é um transgene apresentando extremidades com homologia a alguma região intergênica do genoma (*i.e.*, uma região sem genes) (figura 2).

2.3 Substituição alélica

Adicionalmente, a KI permite outro efeito: a substituição alélica, mencionada no capítulo anterior. Este processo visa a troca da sequência nucleotídica do alelo endógeno por outra sequência exógena ligeiramente diferente, portando SNPs ou pequenas mutações (Zhang et al., 2014). O princípio geral é o mesmo: a administração de Cas9, gRNA e um DNA doador (que é, em si, o novo alelo) permite a recombinação homóloga, resultando na substituição alélica.

2.4 Deleção de genes inteiros

Em algumas situações, a *deleção da sequência completa do gene* é mais interessante do que gerar mutações pontuais nele (por NHEJ ou HDR). Isto porque muitas mutações de ponto eventualmente geram alelos **não nulos**, *i.e.*, que permitem uma atividade ainda que parcial do gene-alvo. Entretanto, a deleção completa de um gene resultará inexoravelmente em uma mutação nula. Uma célula cujos dois alelos do “gene A” foram completamente deletados apresenta um genótipo Δ “gene A”/ Δ “gene A” (Δ é símbolo genético para deleção).

A deleção de genes inteiros pela técnica de CRISPR pode ocorrer tanto por *uma única* quebra da dupla fita de DNA, quanto por *duas* quebras da dupla fita (Sander e Joung, 2014; Mei et al. 2016).

Na primeira, o mecanismo de reparo será sempre HDR, sendo similar ao KO e KI induzidos por recombinação homóloga. Neste caso, porém, vale ressaltar que o DNA doador é composto por uma sequência sem o gene de estudo, apenas com as extremidades homólogas ao cromossomo (figura 3A).

Na segunda, diferentemente das edições gênicas anteriormente descritas, a deleção ocorre por intermédio de *duas* quebras na dupla fita de DNA. Para isso, empregam-se duas endonucleases Cas9 e dois gRNAs com sequências distintas. Na ausência de uma sequência doadora ocorrerá NHEJ, via junção de extremidades que não envolvem o gene-alvo (figura 3B). Contudo, na presença de um DNA doador (de mesma natureza que no exemplo A), ocorrerá o HDR.

3. Marcação de DNA

A técnica de *DNA labeling* associada à CRISPR consiste na marcação *in situ* de locos gênicos com fluorocromos, que possibilitam a visualização tridimensional do loco em células fixadas e também vivas (Anton et al., 2014) sem a necessidade de desnaturação do DNA (Ma et. al, 2015). Técnicas que possibilitam a localização espacial de locos são importantes

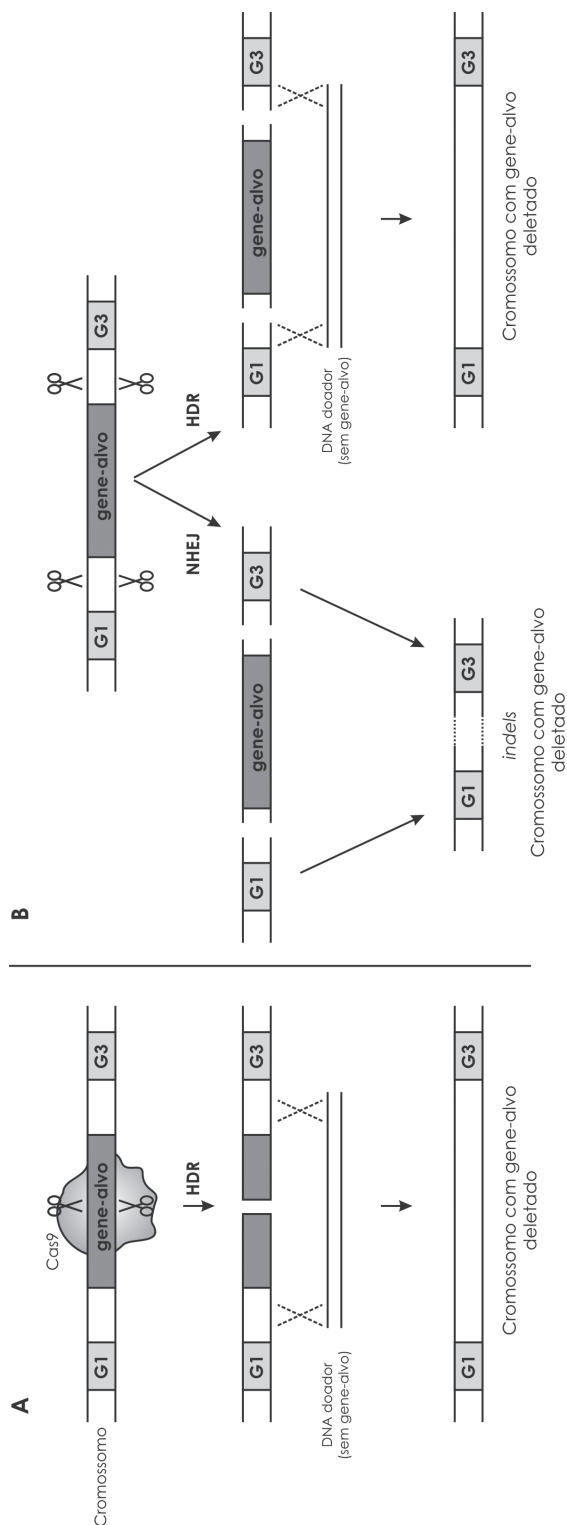


Figura 3. Deleção de genes inteiros. (A) Uma única quebra da dupla fita de DNA, sob o reparo por HDR, leva à deleção do gene inteiro se o DNA doador for composto por uma sequência com homologia com o cromossomo, mas sem o gene-alvo. (B) Quebras duplas na dupla fita de DNA (mediada por dois gRNAs distintos) também podem levar à deleção do gene, por duas maneiras distintas. À esquerda, via NHEJ, a deleção ocorre devido à simples junção de extremidades sem o envolvimento do fragmento do gene-alvo. À direita, via HDR, ocorre o reparo por recombinação homóloga, tendo um DNA doador (o mesmo do exemplo “A”) como molde para a junção. G1 e G3: “gene 1” e “gene 3”, para referência. Imagem: Pereira, TC (2016).

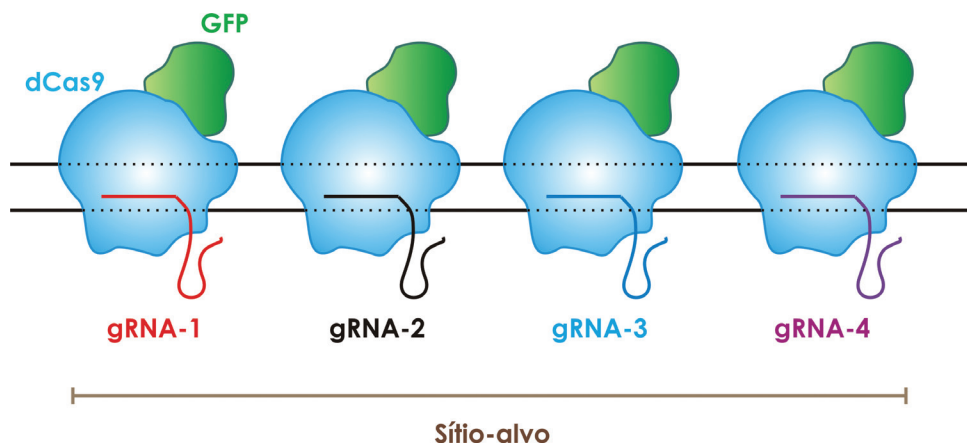


Figura 4. DNA labeling por dCas9. A hibridização *in situ* de locos gênicos ocorre pela ligação de diversas *dead* Cas9 à dupla fita de DNA. gRNA com especificidade para diferentes sítios de um mesmo loco direcionam o local alvo de ligação para dCas9. O acúmulo de dCas9, em associação com proteínas fluorescentes, leva ao aumento pontual de fluorescência, detectável por microscopia de fluorescência de alta resolução. Imagem: Pereira, TC (2016).

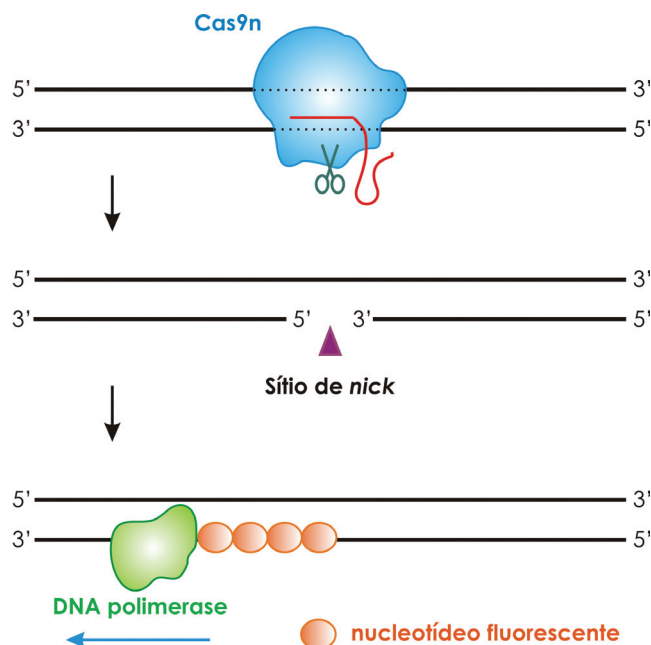


Figura 5. DNA labeling por Cas9n. A marcação de regiões específicas do genoma por nucleotídeos fluorescentes ocorre por intermédio de uma Cas9 modificada (Cas9n), que realiza a quebra apenas na fita 3' -5' do DNA. Em seguida, pela ação da Taq DNA polimerase, nucleotídeos fluorescentes são adicionados adjacentes à quebra da fita dupla, marcando portanto a região específica para o gRNA. Imagem: Pereira, TC (2016).

para o entendimento dos mecanismos de regulação gênica espaço-temporal. Adicionalmente, consiste em uma poderosa ferramenta de diagnóstico genético *in situ*.

Por diversas décadas, técnicas de hibridização *in situ* associadas a ácidos nucleicos fluorescentes (FISH) permitiram a localização de locos gênicos. Porém, devido à necessidade de fixação da célula e desnaturação do DNA, nunca foi possível precisar tais informações em células vivas (Ma et al., 2015).

A presente metodologia consiste na utilização de uma nuclease Cas9 cataliticamente inativa, ou seja, incapaz de clivar a dupla fita de DNA mesmo estando associada gRNA. Devido a esta característica, ela é denominada **dead Cas9 (dCas9)**. Ao gene da dCas9 associa-se uma sequência codificadora de uma proteína de marcação, tal como a proteína verde fluorescente (GFP), proteína vermelha fluorescente (RFP), dentre outras. Portanto, ao ser expressa, a dCas9 apresenta-se indiretamente fluorescente.

Por intermédio da associação de dCas9 a diversos gRNAs específicos para diferentes sítios do mesmo loco, há uma concentração de dCas9 fluorescentes na região de interesse do genoma (figura 4). Consequentemente, por meio de microscopia de fluorescência de alta resolução é possível observar pontos correspondentes ao loco gênico.

Com o avanço contínuo das diversas aplicações do sistema CRISPR nas ciências biológicas, hoje já existem variações da técnica de *DNA labeling*. Dentre as mais relevantes estão a fusão de proteínas fluorescentes, tanto com a dCas9 quanto com o gRNA, desta forma disponibilizando um painel de cores diferentes, que devem se sobrepor no momento de uma ligação dCas9/gRNA, assim como reduzir o *background* de ligações inespecíficas. Também como variáveis da técnica ressalta-se tanto a transfecção dos plasmídeos de expressão de dCas9 e gRNA em células vivas, quanto a metodologia de marcação de células fixadas (Deng et al., 2015). Finalmente, uma metodologia inovadora propõe a adição, por meio de uma DNA polimerase, de nucleotídeos fluorescentes adjacentes ao sítio de corte de fita simples pela Cas9 nickase (Cas9n), desta forma amplificando o sinal de fluorescência e possibilitando uma análise com maior cobertura do genoma (McCaffrey et al., 2015) (figura 5).

4. Regulação da expressão gênica

Uma das primeiras aplicações provenientes da criação dCas9 foi a regulação da expressão gênica (Gilbert et al., 2013). Pesquisadores fusionaram a dCas9 a proteínas reguladoras da transcrição gênica, transformando-a em uma *proteína de transporte*, capaz de levar esses elementos para uma sequência de DNA específica graças ao gRNA. Dependendo da classe de proteína reguladora conduzida pela dCas9, pode-se *induzir* ou *reprimir* a expressão do gene, modificando-se ou não a cromatina ao seu redor. Nesta seção vamos apresentar algumas variantes da dCas9 que podem ser usadas com esse propósito, assim como as vantagens e desvantagens de cada abordagem.

4.1 Para indução da expressão gênica

Induzir a expressão de um gene usando a dCas9 é talvez uma das aplicações mais desafiadoras desse sistema. Isso porque o posicionamento incorreto da dCas9 no promotor ou em suas sequências regulatórias (*enhancers*) pode impedir a ligação de fatores de transcrição endógenos e do complexo da própria RNA polimerase ao DNA. Além disso, altos níveis

de compactação da cromatina, fenômeno típico de genes silenciados, pode dificultar o posicionamento do gRNA.

A versão da dCas9 mais utilizada para indução de transcrição é a dCas9-VP64 (Maeder et al., 2013). Nessa variante, a dCas9 é fusionada a um *tetrâmero* da proteína VP16* (figura 6A), um fator de transcrição do *Herpesvirus simplex* tipo 1 usado para transcrição robusta e imediata de genes no começo da infecção. Embora a dCas9-VP64 seja suficiente para indução de transcrição da maior parte dos genes, também é possível utilizar a versão com 10 domínios do VP16 (dCas9-VP160) caso a indução seja insuficiente. Da mesma maneira, a dCas9-VP48 (com 3 domínios do VP16) pode ser usada para uma indução mais modesta do gene de interesse (Cheng et al., 2013).

Outro aspecto a ser observado para indução de transcrição é a quantidade e o posicionamento dos gRNAs. Salvo raras exceções, múltiplos gRNAs devem ser usados simultaneamente para que qualquer aumento de transcrição seja detectado (Maeder et al., 2013). Geralmente, 3 ou 4 gRNAs posicionados na sequência a montante (extremidade 5') do sítio de início de transcrição (do inglês, *transcription start site*, ou **TSS**) são suficientes para gerar transcrição. Entretanto, em alguns casos 10 gRNAs ou mais podem ser necessários.

Os gRNAs devem ser desenhados de forma que se liguem a sequências próximas ao promotor do gene de interesse. Contudo, os gRNAs devem evitar as regiões utilizadas para a ligação do complexo da RNA polimerase. Os gRNAs mais bem-sucedidos são complementares à fita de DNA que servirá de molde para o RNA, com algumas exceções. Dessa forma, a posição do gRNA vai variar de gene para gene, e cabe ao pesquisador encontrar aqueles adequados para o seu experimento. Por exemplo, para ativar o gene *VEGFA* é preciso colocar os gRNAs entre 450 e 600 pares de bases a montante do TSS. Já para o gene *NFT3*, os melhores gRNAs estão a 135 pares de bases do TSS (Maeder et al., 2013).

Em casos em que a dCas9-VP64 não é suficiente para induzir transcrição, alternativas como a dCas9-SunTag_{10x}-VP64 (figura 6B) podem ser utilizadas (Tanenbaum et al., 2014). O sistema *SunTag* foi criado em 2014 para a amplificação de sinais de fluorescência emitidos por uma proteína. A SunTag é uma sequência de peptídeos que pode ser adicionada à proteína de interesse. Cada peptídeo da SunTag é reconhecido por um fragmento de anticorpo de cadeia simples (scFv, do inglês *single-chain variable fragment*) fusionado a outras proteínas. No caso da dCas9-SunTag_{10x}-VP64, a dCas9 apresenta 10 cópias do peptídeo, que são reconhecidos por 10 anticorpos scFv fundidos a uma VP64 cada. Desta forma, para cada dCas9-SunTag_{10x} presentes no DNA, 10 fatores de transcrição são recrutados. O resultado é um aumento robusto na transcrição do gene de interesse. Em vários casos é possível inclusive induzir um aumento de 3 a 40 vezes na transcrição de um gene com um único gRNA. A desvantagem dessa técnica é que a célula precisa apresentar não somente o gRNA e a dCas9-SunTag_{10x}, mas também o anticorpo scFv contra a sequência de peptídeos. A sua maior vantagem está na possibilidade de se usar somente um gRNA, em vez de 3 ou até 10 gRNAs.

Uma estratégia diferente para indução de transcrição baseia-se em modificar a cromatina onde o TSS se situa. Genes silenciados possuem uma assinatura heterocromática, que inclui metilação e deacetilação de histonas associadas. Por sua vez, genes ativos são caracterizados por pouca metilação e muita acetilação. Por essa razão, pesquisadores começaram a fusionar a dCas9 a enzimas que induzem acetilação de histonas. Essa abordagem é especialmente

* **Nota:** Veja que os múltiplos de VP16 são nomeados de acordo com os valores múltiplos correspondentes. Por exemplo, o trímero de VP16 é chamado VP48 (pois $16 \times 3 = 48$); o tetrâmero, VP64; o decâmero, VP160.

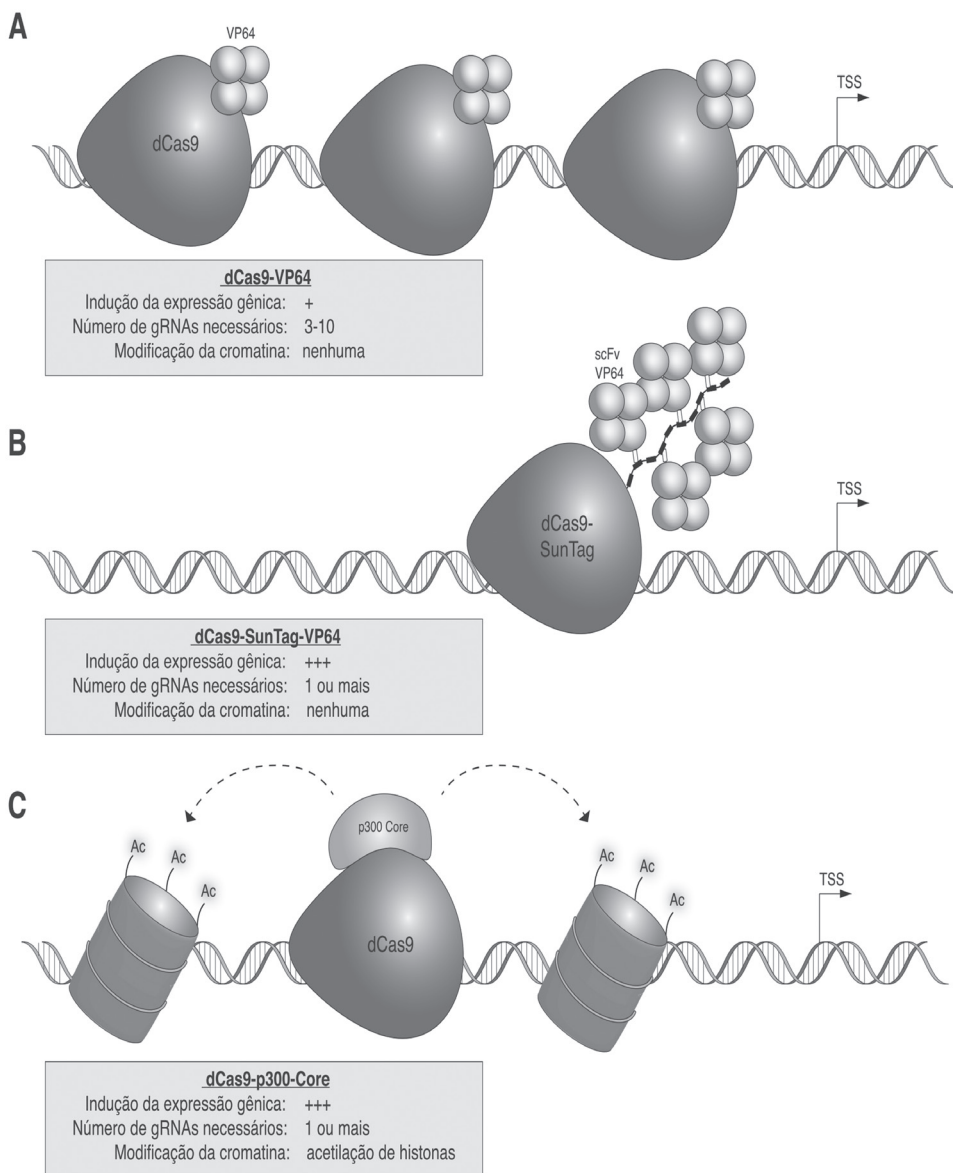


Figura 6. Modelos de uso da dCas9 para indução da expressão gênica. A) dCas9-VP64: a dCas9 é fusionada a 4 cópias do fator de transcrição VP16. Múltiplas cópias da dCas9-VP64 devem ser recrutadas por vários gRNAs às regiões próximas ao promotor do gene de interesse. B) dCas9-SunTag-VP64: a dCas9 é fusionada a uma cadeia de peptídeos (representados como retângulos pretos) e levada ao promotor do gene de interesse por um ou mais gRNAs. A cadeia de peptídeos é reconhecida por anticorpos de cadeia simples fundidos ao VP64. C) dCas9-p300core: variante que, através do núcleo catalítico do p300, induz acetilação de histonas e abertura da cromatina próxima ao TSS (sítio de iniciação da transcrição) do gene de interesse. Imagem: Nunes, PR (2016).

interessante para ativação transcricional de genes presentes em regiões de cromatina muito condensada, onde até a ligação de diversos gRNAs pode ser difícil.

Uma das enzimas acetiltransferases mais conservadas em eucariotos é a p300. O núcleo da p300 (p300^{core}) pode ser fusionado à dCas9 e levada, por meio de um gRNA, até a região regulatória (*enhancer*) do gene de interesse (figura 6C). Uma vez posicionada, o p300^{core} adiciona grupos acetil às histonas vizinhas (principalmente na histona H3, lisina 27) e assim induz aberturas na heterocromatina e o recrutamento de fatores de transcrição (Tsankov et al., 2015). Além disso, o p300^{core} também é capaz de recrutar a RNA polimerase II para formar o complexo de iniciação de transcrição (Cho et al., 1998).

A dCas9-p300^{core} induz transcrição em níveis maiores do que a dCas9-VP64 e, assim como a dCas9-SunTag_{10x}-VP64, pode ser usada com um único gRNA. A vantagem da dCas9-p300^{core} sobre a versão SunTag é que a célula só precisa apresentar a dCas9-p300^{core} e o gRNA, sem necessidade adicional dos anticorpos scFv contra a cadeia de peptídeos.

Continuamente, sistemas cada vez mais complexos e eficientes para indução da expressão gênica podem ser encontrados. A dCas9 foi recentemente fundida a três ativadores de transcrição ao mesmo tempo, formando a dCas9-VP64-p65-Rta. A nova versão promete uma indução 40 a 100 vezes mais forte do que a dCas9-VP64 original (Chavez et al., 2015). Ao mesmo tempo, o próprio gRNA começa a ser modificado para formar estruturas que podem recrutar indiretamente fatores de transcrição. O recrutamento é feito por meio de outras proteínas (fundidas aos fatores de transcrição de interesse) que reconhecem essas estruturas formadas pelo RNA (Konermann et al., 2014).

A indução de transcrição gênica a partir dos próprios promotores endógenos dos genes é uma das aplicações mais poderosas da dCas9. Isto porque diferentes isoformas do transcrito, geradas pelo processo de *splicing*, não podem ser reproduzidas com outras técnicas de superexpressão do DNA complementar (cDNA). Além disso, tornou-se mais fácil e fisiológico induzir a expressão de RNAs longos e não codificantes (lncRNAs) e de RNAs que possuem diferentes TSS. A expressão de múltiplos gRNAs ao mesmo tempo possibilita ainda a indução de vários genes simultaneamente na mesma célula, no processo chamado *multiplexing* (Cheng et al., 2013). Através do uso das variantes da dCas9, também pode-se atingir diferentes níveis de indução de expressão gênica. Cabe ao pesquisador determinar os melhores gRNAs a serem utilizados e a dCas9 mais apropriada para o seu objetivo.

4.2 Para repressão da expressão gênica

Uma das melhores maneiras de se determinar a função de uma proteína ou de um RNA é basicamente eliminá-los da célula e observar quais processos são afetados por sua ausência. Por muitos anos, a técnica de interferência por RNA (RNAi) foi usada com esse propósito. Entretanto, RNAi é um método potencialmente caro e pouco efetivo contra certos alvos, principalmente lncRNAs. Para contornar esse problema, a técnica de *Interferência via CRISPR (CRISPRi)* foi introduzida (figura 7).

Usar a dCas9 para reprimir a expressão de um gene é relativamente mais fácil do que usá-la para induzir transcrição. Isso porque um único gRNA bem posicionado já é suficiente para supressão de até 80% da transcrição (Gilbert et al., 2013). Entretanto, se o uso de somente um gRNA não alcançar a repressão gênica desejada, múltiplos gRNAs podem ser usados.

A dCas9 pode ser utilizada sozinha ou em sua forma fusionada ao repressor de transcrição KRAB (do inglês, *Krüppel-associated box domain*). Em ambos os casos, a dCas9

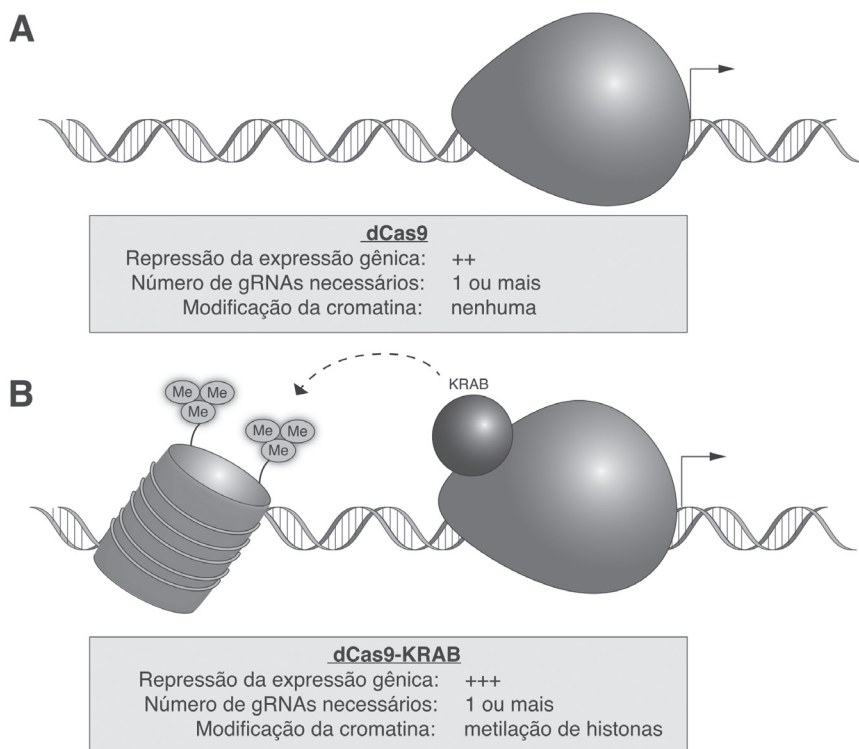


Figura 7. Modelos de uso da dCas9 para repressão da expressão gênica. A) dCas9: esta proteína sozinha pode ser usada para impedir a ligação de fatores de transcrição e do complexo da RNA polimerase II ao DNA. Para isso, deve-se posicioná-la perto do TSS (sítio de iniciação de transcrição) do gene de interesse. B) dCas9-KRAB: essa variante, além de servir de barreira física contra a transcrição, também modifica a cromatina próxima do TSS de maneira a torná-la mais heterocromática e menos acessível à maquinaria de transcrição. Imagem: Nunes, PR (2016).

deve ser direcionada ao promotor de um gene ou ao seu TSS. A dCas9 liga-se com grande afinidade ao DNA através do gRNA e bloqueia a ligação de fatores de transcrição e do complexo da própria RNA polimerase, impedindo a transcrição (Gilbert et al., 2013).

Já a dCas9-KRAB, além de impedir fisicamente a ligação de outras proteínas ao DNA, também promove a formação de heterocromatina ao seu redor. O domínio KRAB recruta indiretamente proteínas da família da HP1 (do inglês, *heterochromatin protein 1*) que, por sua vez, interagem com histonas metiltransferases, induzindo metilação da histona H3, por exemplo.

O local de ancoragem (posicionamento) da dCas9 (com ou sem KRAB) também é um aspecto essencial para repressão eficiente da transcrição. O gRNA deve ser dirigido à região do TSS, ou até 100 pares de bases antes ou depois do mesmo (Gilbert et al., 2013). Em geral, quanto mais longe do TSS estiver a dCas9, menos efetiva a repressão será. A exceção acontece quando a dCas9-KRAB é posicionada de forma a silenciar *enhancers* distantes (a dezenas de milhares de pares de bases do TSS) (Thakore et al., 2015). Tal estratégia pode ser usada para mapear *enhancers* de genes específicos, mas a repressão em si é menos robusta do que aquela obtida com posicionamento da dCas9-KRAB próxima ao TSS.

Assim como na indução da expressão gênica, múltiplos gRNAs podem ser usados simultaneamente em combinação com a dCas9 ou dCas9-KRAB de forma a reprimir a expressão de vários genes ao mesmo tempo (*multiplexing*), o que não é facilmente feito com técnicas de RNAi. A CRISPRi também possibilita a repressão da expressão de proteínas na fonte, por interferir diretamente com a produção do mRNA (*i.e.*, silenciamento gênico transcricional). Da mesma forma, tornou-se mais fácil controlar os níveis de lncRNAs e de outros RNAs contra os quais técnicas tradicionais de RNAi não são eficientes. Embora seja uma técnica simples, rápida e robusta, o pesquisador deve determinar a melhor posição da dCas9 ou da dCas9-KRAB e se mais de um gRNA é necessário.

5. Clivagem de RNAs

O domínio catalítico da Cas9 que corta a fita de DNA complementar ao gRNA (domínio chamado HNH) tem a mesma estrutura de alguns domínios que cortam RNA. A partir dessa observação, a possibilidade de se usar a Cas9 para clivar RNAs começou a ser estudada. O maior problema encontrado é que a Cas9 se liga com pouca afinidade ao RNA de fita simples (ssRNA), e não se liga a níveis detectáveis ao RNA de fita dupla (dsRNA). Para induzir a associação da Cas9 com o RNA é necessário prover a célula com sequências chamadas **PAMmers** (figura 8): cadeias de DNA em fita simples (ssDNA) formadas por 10 ou até 20 nucleotídeos e cujas extremidades 5' contêm a sequência PAM (5'-NGG-3', ou somente 5'-GG-3') (O'Connell et al., 2014). A função dos PAMmers é se parear com o RNA-alvo e apresentar a sequência PAM à Cas9. Isso induz uma mudança de conformação da Cas9 e permite que seus sítios catalíticos clivem o RNA.

Essa estratégia leva à clivagem eficiente do RNA de interesse, mas também pode levar à associação do complexo Cas9/gRNA com outros RNAs contendo PAMmers. Para reduzir a associação da Cas9/gRNA com alvos indesejados, os PAMmers devem ser estendidos em até 8 bases na extremidade 5' (figura 8B). Nesse caso, parte do PAMmer fica alinhado à sequência reconhecida pelo gRNA, impedindo a ligação do gRNA ao RNA sem que haja dissociação de parte do complexo PAMmer-RNA. Essa dissociação necessita da energia que é normalmente liberada pelo pareamento de bases entre o gRNA e o RNA-alvo. Assim, a ligação do gRNA a um RNA que não tem uma sequência compatível com o gRNA torna-se energeticamente desfavorável, reduzindo a afinidade do complexo Cas9/gRNA por outros RNAs.

Outro problema em potencial é a provável clivagem do DNA que deu origem ao RNA-alvo pela Cas9, já que ambos possuem a mesma sequência. Para contornar este problema, os PAMmers devem se direcionar à parte do RNA onde não haja uma sequência PAM presente, criando um desalinhamento de três pares de base (figura 8C). A sequência PAM presente no PAMmer, mesmo não complementar ao RNA-alvo, é suficiente para recrutar e ativar a Cas9, enquanto a ausência da PAM no genoma impede que ele seja clivado pela Cas9 (O'Connell et al., 2014). Tal estratégia também permite a clivagem de RNAs que eventualmente não possuam uma sequência PAM.

Em resumo, PAMmers simples (figura 8A) são os mais eficientes em induzir clivagem do RNA-alvo, porém existe a possibilidade de clivagem de outros RNAs contendo PAMmers e do DNA. PAMmers com extensão na extremidade 5' (figura 8B) reduzem significativamente a clivagem de outros RNAs, mas também podem diminuir em até 10 vezes a clivagem do RNA-alvo. PAMmers com sequência PAM não complementar ao RNA-alvo (figura 8C) eliminam a

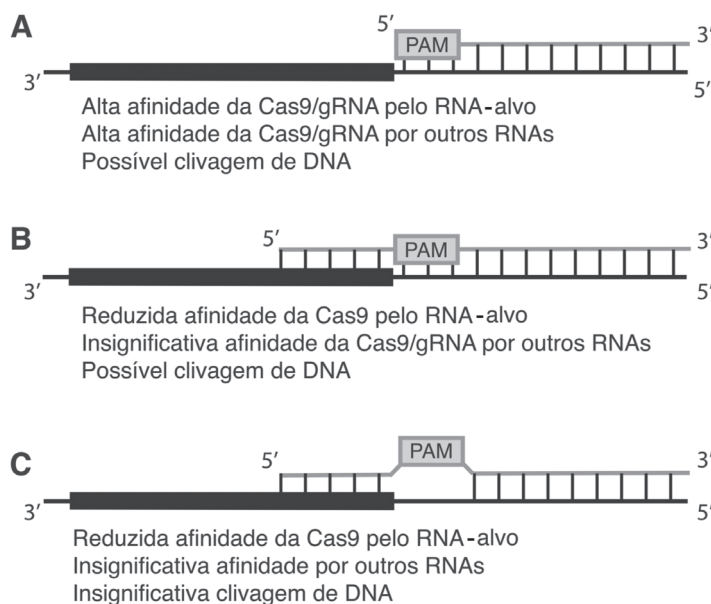


Figura 8. Clivagem de RNA pela Cas9 e os diferentes PAMmers que podem ser utilizados. O RNA é mostrado em preto (linha inferior) e o retângulo preto representa a sequência do RNA reconhecida pelo gRNA. Os PAMmers (ssDNA) são mostrados em cinza (linha superior). A) PAMmer simples, com a sequência PAM no seu final 5'. Este PAMmer tem alta afinidade pelo RNA-alvo, mas também se liga a outros RNAs e ao DNA. B) O PAMmer com uma extensão na sua extremidade 5' tem maior afinidade pelo RNA-alvo e menor afinidade por outros RNAs. Entretanto, esses PAMmers dificultam a interação do complexo Cas9/gRNA com o RNA, já que a ligação entre o PAMmer e o RNA-alvo deve ser desfeita antes do alinhamento do gRNA com o RNA de interesse. C) O PAMmer com uma sequência PAM não complementar ao RNA-alvo tem menor afinidade pelo mesmo, mas não leva à clivagem do DNA que deu origem ao RNA de interesse. Imagem: Nunes, PR (2016).

clivagem do DNA. Todas as modificações do PAMmers têm vantagens e desvantagens. Cabe ao pesquisador escolher a abordagem de acordo com a eficiência e especificidade desejadas.

A mesma técnica pode ser usada também para precipitação de RNAs específicos. Para isso, deve-se usar a *dead* Cas9 em vez da Cas9 cataliticamente ativa e utilizar PAMmers feitos de oligômeros modificados com 2' O-Metil, o que impede que RNases degradem a dúplice RNA:DNA (O'Connell et al., 2014). A dCas9 nesse caso pode ser precipitada com um anticorpo contra a própria dCas9 ou contra uma *tag* adicionada a ela.

O uso da Cas9/dCas9 para clivar ou precipitar RNAs específicos pode transformar o modo como estudamos estas moléculas. A técnica promete a clivagem específica dos RNAs selecionados e a possibilidade de se isolar um RNA sem que uma *tag* tenha sido inserida nele previamente. Além disso, seu baixo custo e facilidade interessa toda a comunidade científica.

6. Mapeamento de genes

Um dos principais objetivos da genética moderna consiste na localização e identificação de sequências de DNA responsáveis por determinadas variações fenotípicas. Em qual(ais) cromossomo(s) estão os genes responsáveis pela cor dos olhos? E pelo tamanho das asas em

moscas de fruta? E por certas deficiências como hemofilia ou talassemia? Algumas destas perguntas puderam ser respondidas através dos estudos iniciais de Morgan e de seu aluno Sturtevant, como a herança de genes localizados no mesmo cromossomo. Nesses estudos, a taxa de recombinação gênica meiótica entre dois genes era correspondente à distância destes no cromossomo (Snustad e Simmons, 2012).

A partir disso, o mapeamento genético vem sendo realizado através do estudo da cosegregação (herança conjunta) de determinados marcadores genéticos com o fenótipo de interesse em um amplo grupo de indivíduos, para assim determinar as regiões específicas do genoma que influenciam este fenótipo (Snustad e Simmons, 2012). Entretanto, esta forma de localização depende da taxa de recombinação meiótica na região entre os diferentes marcadores genéticos que, em geral, é muito baixa. Dessa forma, na prática, a localização específica de genes ou variantes internas de cada gene, como polimorfismos, torna-se inviável, visto que exigem construções de painéis imensos de indivíduos ou estudos de várias gerações.

Um novo método também baseado em eventos de recombinação direcionados a regiões de interesse vem sendo desenvolvido através do estudo de recombinações mitóticas, permitindo o aumento da resolução do mapeamento (Yin e Petes, 2013). A recombinação mitótica ocorre de forma rara durante o reparo de quebras de fita dupla por recombinação homóloga, gerando um cromossomo recombinante. Supondo uma célula heterozigótica, este processo, conhecido por perda da heterozigosidade, resultaria em células filhas cujo genótipo é completamente homozigoto do ponto de quebra e reparo até o telômero; no restante do cromossomo, ele permanece semelhante à célula-mãe – heterozigoto. Assim, para construção do mapa genômico utilizam-se indivíduos com perda da heterozigosidade em diferentes pontos do genoma associados com a variação do fenótipo estudado (Henson et al., 1991). Na prática, estes estudos também são limitados devido à baixa frequência em que esses eventos de recombinação mitótica ocorrem naturalmente.

É neste contexto que o sistema CRISPR-Cas9 pode ser aplicado no processo de mapeamento genético. Esse sistema facilita a construção de painéis de indivíduos com diferentes níveis de perda de heterozigosidade através da indução de eventos de recombinação mitótica em locais específicos (Sadhu et al., 2016) (figura 9). Como já elucidado no capítulo 2, a endonuclease Cas9 é capaz de causar quebra dupla na fita de DNA em uma sequência específica de acordo com o RNA-guia ligado a ela. No entanto, para que Cas9 se ligue e clive o DNA com sucesso é necessária uma sequência PAM localizada na fita complementar à sequência-alvo (Doudna e Charpentier, 2014).

Para a ocorrência da perda da heterozigosidade em indivíduos heterozigotos como explicado acima, é necessário a quebra de dupla fita em apenas um dos cromossomos homólogos. Mas como garantir que CRISPR atue em apenas um dos cromossomos? Este objetivo é alcançado com a utilização de sítios PAM polimórficos, garantindo que apenas um dos cromossomos seja afetado pela quebra da fita dupla (Sadhu et al., 2016).

Recentemente, a aplicação de CRISPR permitiu mapear a sensibilidade a manganês a um polimorfismo único no transportador Pmr1 (Sadhu et al., 2016), mostrando alta resolução da técnica de mapeamento. Para tanto, inicialmente foram desenvolvidos 95 RNAs-guias que direcionaram Cas9 a sítios distribuídos ao longo de todo o braço esquerdo do cromossomo 7 da levedura *Saccharomyces cerevisiae*. A cepa utilizada no estudo era diploide heterozigótica originária do cruzamento de uma cepa de laboratório (BY) e um cepa de vinhedo (RM). Depois do processo de quebra e reparo, as células que sofreram reparo foram separadas por citometria de fluxo de acordo com a perda do gene GFP próximo ao telômero. O sequenciamento

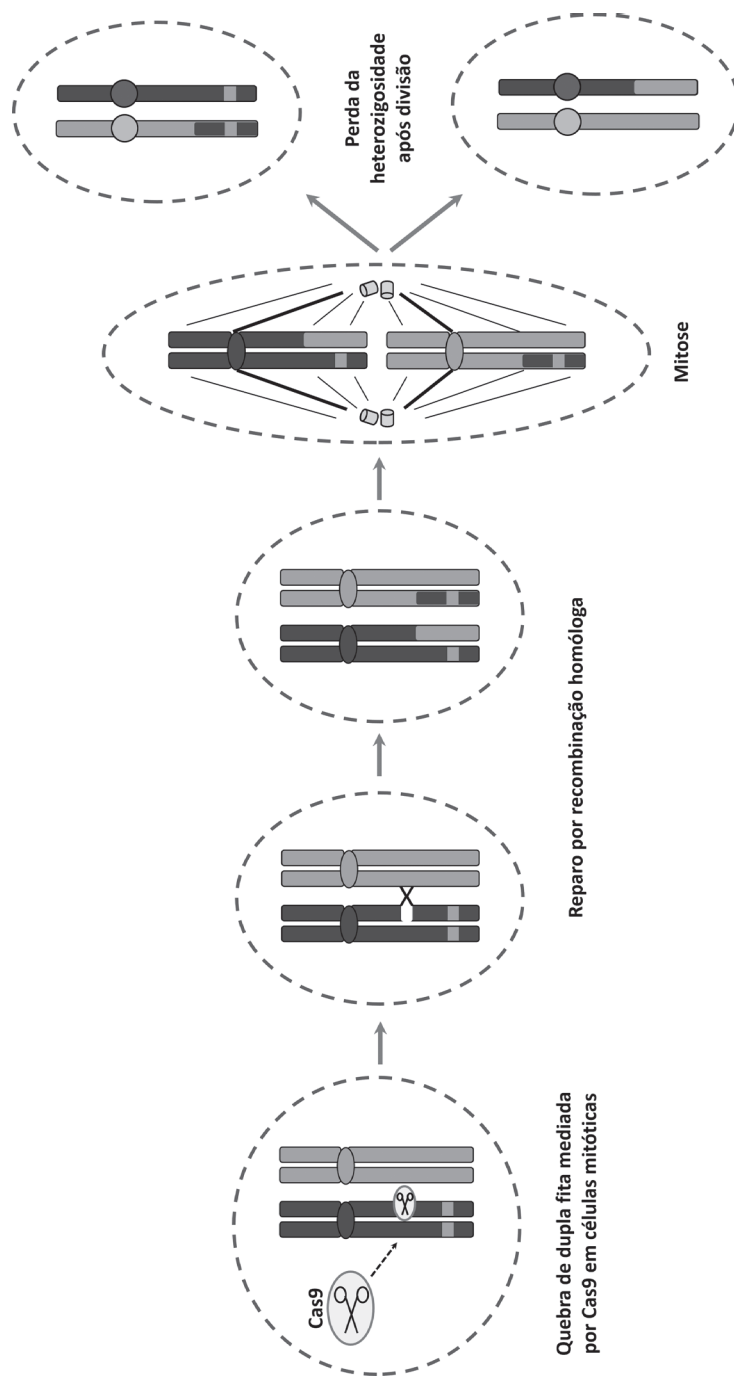


Figura 9. Perda de heteroziguidade induzida por sistema CRISPR-Cas9. Os indivíduos com perda de heteroziguidade (LOH) são isolados pela ausência do marcador dominante (banda preta). O processo pode ser repetido várias vezes e assim, pelo seguimento de marcadores, é possível mapear as variantes gênicas de uma determinada região cromossômica. Adaptado de Sadhu et al., 2016. Imagem: Roberto, GM e Annichini, MSB (2016).

posterior de 4 linhagens originárias de cada RNA-guia mostrou que 75% dos eventos de recombinação ocorreram de forma bem-sucedida em torno de 20 kb da sequência-alvo. O crescimento dessas linhagens foi observado em 12 condições diferentes que já haviam sido relacionadas ao cromossomo 7. A análise estatística mostrou que a perda da heterozigose afetava a sensibilidade ao manganês das linhagens com recombinação após a posição 180-200 kb. Uma análise mais fina avaliando um maior número de cepas cuja ligação e quebra via Cas9 ocorreram especificamente nesta região permitiu uma identificação mais precisa do intervalo em que as variantes se separam, e assim a descoberta de um polimorfismo responsável por aumentar a sensibilidade nas cepas BY ao manganês. Foi visto que esta variante codifica uma fenilalanina em BY e uma leucina de RM no gene *Pmr1*, um transportador de manganês.

O mapeamento por perda de heterozigose já se mostrava interessante, pois o restante do genoma permanece intacto, aumentando a confiabilidade do resultado. Com a introdução de CRISPR à técnica, a especificidade desta aumenta, pois torna-se possível gerar alta frequência de eventos de recombinação e acessar qualquer região do genoma, tudo isso em pouco tempo. Dessa forma, essa metodologia facilita o entendimento da genética por trás das variações fenotípicas (Sadhu et al., 2016).

Além disso, essa metodologia pode ser aplicada não somente para mapeamento mas também no caso de certos alelos relacionados a doenças cujo acesso direto por CRISPR é difícil devido à complexidade da sequência. Estes alelos podem ser reparados pela indução de DSB nas proximidades do alelo patogênico e sua substituição pelo alelo sadio através de perda de heterozigosidade (Sadhu et al., 2016).

7. Rastreamento de RNAs

Embora atualmente existam diversas e bem estabelecidas maneiras de reduzir a expressão gênica através do emprego de RNAs de interferência ou de sequências nucleotídicas antissenso, um estudo mais profundo relacionado à regulação pós-transcricional dos RNAs mensageiros como o tráfico intracelular, formas de *splicing* e poliadenilação, bem como os parâmetros espaço-temporais envolvidos na tradução, ainda carecem de uma metodologia mais robusta para avaliação.

A localização do RNA na célula, bem como o momento e a forma pela qual ele chega a um determinado compartimento subcelular, podem influenciar se as proteínas correspondentes estão sendo produzidas no local ou na hora correta. Diversas doenças são decorrentes diretamente de alterações no RNA; defeitos no transporte de mRNAs ou nos mecanismos de *splicing* estão relacionados a quadros como autismo ou câncer, por exemplo (Srebrow e Kornblihtt, 2006; Liu-Yesucevitz et al., 2011). Dessa forma, a elucidação destes processos se beneficiaria imensamente de técnicas para rastreamento de RNAs em células vivas, permitindo observar desde sua localização até suas variações ao longo do eixo transcrição-tradução.

Algumas técnicas têm sido empregadas para rastrear moléculas de RNA utilizando proteínas, oligonucleotídeos complementares ao alvo ou uma combinação dos dois (Lu et al., 2009; Sokol et al., 1998; Fouts et al., 1997; Bertrand et al., 1998). Entretanto, todas elas necessitam de sofisticada manipulação genética e não servem para reconhecimento arbitrário de sequências de RNA, além de potencialmente influenciarem o papel do RNA endógeno.

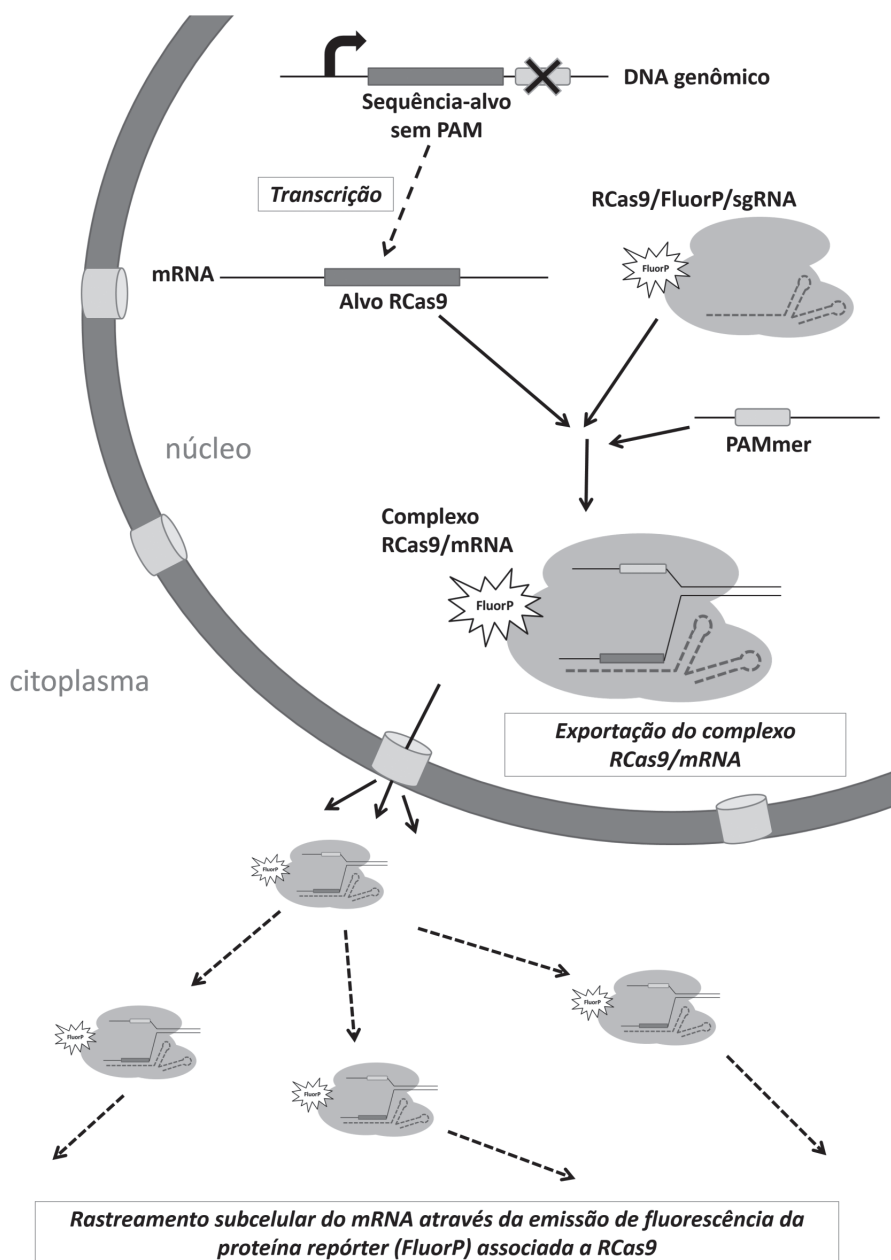


Figura 10. Rastreamento de RNA em células vivas através de sistema CRISPR/RCas9 associado a uma proteína fluorescente e sequência PAMmer. Imagem: Roberto, GM e Annichini, MSB (2016) adaptado de Nelles et al. (2016).

Nelles e colaboradores (2016) mostraram que a ligação de dCas9:GFP, associada a um gRNA, é capaz de se ligar à sequência 3' UTR do mRNA-alvo ainda no núcleo (sistema denominado de **RCas9**) sem interferir com sua exportação para o citoplasma, sua abundância ou a quantidade de proteína produzida a partir dele (Nelles et al., 2016). Esta abordagem

apresenta certas vantagens em relação às outras técnicas de rastreamento de RNAs, visto que o reconhecimento se baseia apenas na especificidade do gRNA, sem a necessidade de manipulação genética e outras proteínas ligantes de RNA, permitindo expandir as habilidades dos pesquisadores para rastrear os transcritos.

Dentre as futuras aplicações desta técnica pode-se citar: (i) a avaliação ou alteração do *splicing* por meio do recrutamento de diferentes proteínas e (ii) o rastreamento de RNAs podendo levar à descoberta de diferenças no padrão de expressão gênica entre indivíduos saudáveis e doentes. A longo prazo torna-se possível a modulação de muitas características e reprogramação do comportamento celular via alterações no processamento do RNA, na expressão gênica ou na concatenação de enzimas no RNA-alvo visando aplicação na biologia sintética e no tratamento de doenças.

8. Conclusões

É admirável observar como, a partir de apenas dois elementos – Cas9 e gRNA –, os pesquisadores têm sido capazes de expandir a caixa de ferramentas moleculares baseadas em CRISPR (figura 11). Apenas poucos anos após o lançamento dos alicerces da técnica, diversas estratégias derivadas de CRISPR foram desenvolvidas e provavelmente muitas outras virão nos próximos anos.

	Ação de Cas9	Alvo	Variação da técnica	Objetivo
CRISPR	Clivagem	DNA	Edição genética (KO; KI; SA e DG)	Modificação
			Mapeamento gênico	Localização
		RNA	Clivagem de RNA	Expressão
	Ligação	DNA	Regulação da expressão gênica	
			Marcação do DNA	Localização
		RNA	Rastreamento de RNA	

Figura 11. Relação entre as formas de ação de Cas9 sobre diferentes ácidos nucleicos e as técnicas derivadas. Imagem: Pereira, TC (2016).

Referências

- Anton T, Bultmann S, Leonhardt H, Markaki Y. Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus*. 2014;5(2):163-72.
- Bertrand E, Chartrand P, Schaefer M, Sheno SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. *Mol Cell*. 1998;2:437-45.
- Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle MPR, Iyer E, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*. 2015;12:326-8.
- Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*. 2013;23:1163-71.

- Cho H, Orphanides G, Sun X, Yang X-J, Ogryzko V, Lees E, et al. A human RNA polymerase II complex containing factors that modify chromatin structure. *Mol Cell Biol*. 1998;18:5355-63.
- Deng W, Shi X, Tjian R, Lionnet T, Singer RH. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc Natl Acad Sci USA*. 2015;112(38):11870-5.
- Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346:1258096. <http://dx.doi.org/10.1126/science.1258096>.
- Fouts DE, True HL, Celandier DW. Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. *Nucleic Acids Res*. 1997;25:4464-73.
- Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. 2013;154:442-51.
- Henson V, Palmer L, Banks S, Nadeau JH, Carlson GA. Loss of heterozygosity and mitotic linkage maps in the mouse. *Proc Natl Acad Sci USA*. 1991;88:6486-90.
- Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157(6):1262-78.
- Klug WS, Cummings MR, Spencer CA, Palladino MA. *Conceitos de Genética*. 9. ed. Artmed; 2010.
- Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2014;517:583-8.
- Lewin, B. *Genes IX*. 9. ed. Artmed; 2008.
- Liu-Yesucevitz L, Bassell GJ, Gitler AD, Hart AC, Klann E, Richter JD, et al. Local RNA translation at the synapse and in disease. *J Neurosci*. 2011;31(45):16086-93.
- Lu G, Dolgner SJ, Tanaka TM. Understanding and engineering RNA sequence specificity of PUF proteins. *Hall Curr Opin Struct Biol*. 2009;19(1):110-5.
- Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S, Pederson T. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci USA*. 2015;112(10):3002-7.
- Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*. 2013;10:977-9.
- McCaffrey J, Sibert J, Zhang B, Zhang Y, Hu W, Riethman H, et al. CRISPR-CAS9 D10A nickase target-specific fluorescent labeling of double strand DNA for whole genome mapping and structural variation analysis. *Nucleic Acids Res*. 2016;29;44(2):e11.
- Mei Y, Wang Y, Chen H, Sun ZS, Ju XD. Recent Progress in CRISPR/Cas9 Technology. *J Genet Genomics*. 2016;43(2):63-75.
- Meru J, Sadhu JS, Bloom LD, Kruglyak L. Crispr-directed mitotic recombination enables genetic mapping without crosses. *Science*. 2016. <http://dx.doi.org/10.1126/science.aaf5124>.
- Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, et al. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell*. 2016;165(2):488-96. <http://dx.doi.org/10.1016/j.cell.2016.02.054>.
- O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*. 2014;516:263-6.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-308.
- Sadhu MJ, Bloom JS, Day L, Kruglyak L. CRISPR-directed mitotic recombination enables genetic mapping without crosses. *Science*. 2016 May;352(6289):1113-6.
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014;32(4):347-55.
- Snustad P, Simmons M. *Principles Of Genetics*. Hoboken: Wiley; 2012.
- Sokol DL, Zhang X, Lu P, Gewirtz AM. Real time detection of DNA. RNA hybridization in living cells. *Proc Natl Acad Sci USA*. 1998;95:11538-43.
- Srebrow A, Kornbliht AR. The connection between splicing and cancer. *J Cell Sci*. 2006;119(Pt 13):2635-41.
- Strachan T, Read A. *Genética Molecular Humana*. 4. ed. Artmed; 2013.
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159:635-46.
- Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*. 2015;12:1143-9.
- Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, Amit I, et al. Transcription factor binding dynamics during human ES cell differentiation. *Nature*. 2015;518:344-9.
- Yin Y, Petes TD. Genome-wide high-resolution mapping of UV-induced mitotic recombination events in *Saccharomyces cerevisiae*. *PLOS Genet*. 2013. <http://dx.doi.org/10.1371/journal.pgen.1003894>.
- Zhang C, Xiao B, Jiang Y, Zhao Y, Li Z, Gao H, et al. Efficient editing of malaria parasite genome using the CRISPR/Cas9 system. *MBio*. 2014 Jul;5(4):e01414-14.

Geraldo Aleixo Passos^{1,2,3}, Cesar Augusto Speck Hernandez^{1,3}, Larissa Cotrim de Sousa¹,
Rafaela de Freitas Martins Felício^{1,3} e Tiago Alves Jorge de Souza^{4,5}

¹*Grupo de Imunogenética Molecular, Dept.º de Genética, FMRP, USP, SP - Brasil*

²*Dept.º de Morfologia, Fisiologia e Patologia Básica, FORP, USP, SP - Brasil*

³*Programa de Pós-Graduação em Imunologia Básica e Aplicada, FMRP, USP, SP - Brasil*

⁴*Dept.º de Genética, FMRP, USP, SP - Brasil*

⁵*Programa de Pós-Graduação em Genética, FMRP - USP, SP - Brasil*

Estruturação do capítulo

1. Introdução
2. Aplicações na pesquisa básica
 - 2.1 Identificação da função gênica
 - 2.2 Construção de modelos para pesquisa (células e organismos)
3. Aplicações na medicina
 - 3.1 Combate a patógenos
 - 3.2 Combate ao câncer
 - 3.3 Terapia genética
 - 3.4 Controle de vetores de doenças
4. Aplicações na agricultura e veterinária
 - 4.1 Combate a pragas agrícolas
 - 4.2 Geração de variação genética – novos cultivares não transgênicos
 - 4.3 Animais resistentes a doenças
5. Aplicações na indústria
 - 5.1 Produção de biomoléculas
 - 5.2 Produção de biocombustíveis
 - 5.3 Produção de biomateriais
6. Conclusões

1. Introdução

Pesquisadores interessados na técnica de CRISPR, mas que ainda não conhecem a amplitude de possibilidades dessa ferramenta, terão neste capítulo a oportunidade de conhecer um pouco melhor sobre algumas das diversas aplicações desta. Como visto na figura 1, faremos um apanhado geral dos usos de CRISPR em diversas áreas, desde a investigação científica básica até as aplicações práticas para a saúde humana, o agronegócio e o setor industrial. Contudo, é importante que o leitor tenha em mente que o número de aplicações cresce continuamente.

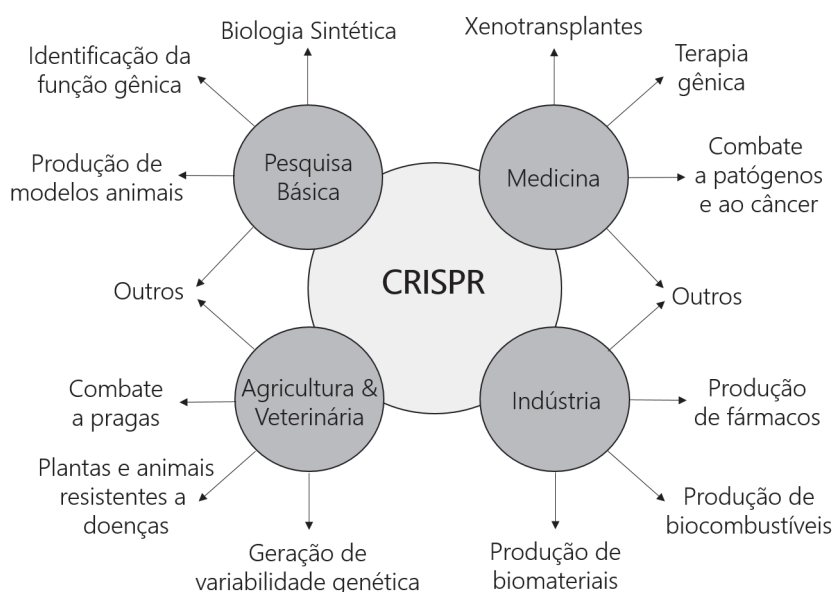


Figura 1. Diversidade de aplicações de CRISPR. A ampla gama de efeitos moleculares (ativação/repressão gênica, substituição alélica, clivagem de DNA/RNA, deleção de genes, *knock-in/out*, inserção de transgenes, geração de variação genética, etc.) e abordagens possíveis (*DNA labeling*, mapeamento gênico, *RNA tracking*, etc.) por meio de CRISPR se traduz em uma miríade de possibilidades para a ciência, tecnologia e mercado. Imagem: Pereira, TC (2016).

2. Aplicações na pesquisa básica

2.1 Identificação da função gênica

Estima-se que existam em torno de 20-25 mil genes no genoma humano (<https://www.encodeproject.org/>), muitos dos quais ainda sem função conhecida. Para evidenciar o papel de um gene em um processo biológico, ou sua relação com determinada doença genética, uma das estratégias mais utilizadas é a *perturbação*, ou mesmo a *anulação*, da expressão deste gene e a posterior observação das consequências que isto provoca (na célula ou no organismo). Essa perturbação genética (*e.g.*, nocauteando ou silenciamento do gene) pode ser realizada por CRISPR, de diversas formas diferentes.

Um exemplo claro, simples e direto desta estratégia foi demonstrado por Ecsedi e colaboradores (2015). Já se sabia que o microRNA *let-7* estava indiretamente associado ao desenvolvimento da vulva no nematoide *C. elegans*, por meio da repressão de algum mRNA-alvo desconhecido. Por meio do nocauteamento dos melhores genes-alvos candidatos, os autores determinaram que o gene *lin-41* era de fato o gene responsável por esta característica em *C. elegans* (Ecsedi et al., 2015).

Contudo, análises mais sofisticadas podem ser aplicadas. Tomemos como exemplo o sistema imune, que depende de um circuito transcricional regulado de maneira precisa e numa escala genômica (Kim et al., 2009). Parnas e colaboradores desenharam uma biblioteca com mais de 125.000 gRNAs (seis gRNAs por gene) dirigidos contra diferentes alvos no genoma de células dendríticas derivadas de camundongos que expressam constitutivamente o gene da Cas9. O uso de lentivírus (1 por célula) permitiu a inserção do gRNA nos genomas das células. Como as células dendríticas respondem ao estímulo induzido por LPS produzindo TNF, foi então possível determinar quais genes são alterados tanto positiva como negativamente sob esse estímulo. Os resultados desse experimento permitiram identificar genes nunca antes relacionados com a expressão de TNF, como *Tti2*, *Ruvbl2*, *Tmem258*, *Midn*, *Ddx39b*, *Stat5b* e *Pdcd10* (Parnas et al., 2015).

CRISPR também tem sido utilizada na identificação de efeitos de certo gene em uma determinada patologia. Neste caso, os pesquisadores demonstraram como a geração de mutações pontuais específicas em oncogenes pode guiar o desenvolvimento de adenocarcinomas. O gene *Kras* foi o escolhido, visto que 33% dos pacientes com adenocarcinoma de pulmão apresentam a mutação G12D nele. Além disso, os autores geraram mutações do tipo *indels* nos genes *p53* e *Lkb1*, que são amplamente relacionados ao desenvolvimento desse tipo de câncer. Para isso, vírus adenoassociados portando gRNAs foram administrados *in vivo* em camundongos, posteriormente os genomas de células dos pulmões desses animais foram então sequenciados. Os pesquisadores observaram que os genes *p53* e *Lkb1* apresentaram *indels* (0,1% no gene *p53* e 0,4% no gene *Lkb1*), sendo também possível ditar o gene *Kras* (0,1% de eventos de recombinação). Passados dois meses, os camundongos modificados apresentaram tumores e nódulos cancerígenos no pulmão, com uma heterogeneidade alta no espectro das mutações nas células tumorais. Este resultado é muito interessante, já que este tipo de heterogeneidade também acontece na clínica, ou seja, pacientes diferentes podem apresentar alterações genéticas distintas considerando um mesmo tipo de tumor (Platt et al., 2014).

2.2 Construção de modelos para pesquisa (células e organismos)

A criação de modelos celulares para determinadas disfunções é essencial para a compreensão dos mecanismos biológicos subjacentes. Semelhantemente, a criação de animais modelos de doenças genéticas permite não apenas o entendimento da fisiopatologia destas, mas também o desenvolvimento e teste de possíveis tratamentos. A seguir veremos alguns exemplos.

Modelos celulares

A linhagem celular Madin-Darby, derivada de rim de cão (*Canis familiaris*), é comumente utilizada em estudos de transporte de drogas. O objetivo de Simoff e colaboradores (2016) foi gerar uma linhagem celular Madin-Darby nocaute (KO) com uma disfunção na produção da

proteína Mdr1, parte da superfamília de transportadores ABC (*ATP-binding cassette*), grupo de proteínas envolvidas no transporte e excreção de drogas. Para isto eles geraram dois vetores CRISPR específicos para os éxons 4 e 13 do gene *Mdr1* canino. Depois de selecionar um clone celular e mantê-lo em cultura, foi possível verificar a ausência de produção da proteína e uma alteração nas vias de transporte mediadas pela Mdr1, estabelecendo assim um novo modelo celular de estudo para o transporte de drogas (Simoff et al., 2016).

Modelos animais

Armstrong e colegas utilizaram CRISPR para inserir mutações no genoma de *zebrafish* (*Danio rerio*), aproveitando o sistema de HDR da célula e utilizando um molde doador de fita simples de DNA (ssODN) para obter uma edição específica. O objetivo dos autores era gerar linhagens de animais *knock-in* para os genes *tardbp* e *fus*, os quais codificam duas proteínas relacionadas ao desenvolvimento de esclerose lateral amiotrófica (ALS). Até então, dois polimorfismos haviam sido descritos como de risco para o desenvolvimento de ALS em humanos (*TARDBP*^{A382T} e *FUS*^{R521H}). Neste sentido, os autores injetaram nos embriões de *zebrafish* um vetor CRISPR junto com um molde de DNA homólogo contendo a sequência com as mutações de interesse. Os pesquisadores reportaram substituições específicas no genoma com uma eficiência de 30% no gene *tardbp* e 26% no gene *fus*, obtendo assim novos modelos píceos para ALS (Armstrong et al., 2016).

Outro exemplo interessante é o emprego de CRISPR para gerar coelhos (*Oryctolagus cuniculus*) nocauteados para o gene *GJA8* (que codifica uma *gap junction protein*), cuja função é crucial para o desenvolvimento normal das lentes nos globos oculares nos vertebrados. Em humanos, as alterações nesse gene podem desencadear cataratas congênitas em crianças. Os pesquisadores desenharam dois vetores contendo sgrNAs dirigidos contra a sequência consenso do gene *GJA8*. Os vetores foram microinjetados em zigotos de coelhos e transferidos aos ovidutos das fêmeas, gerando assim coelhos deficientes desse gene. Os resultados do estudo mostraram que a alteração na expressão do gene *GJA8* causou microftalmia, lentes de menor tamanho e cataratas, demonstrando assim a importância deste no desenvolvimento ocular (Yuan et al., 2016).

Suínos domésticos (*Sus scrofa domesticus*) também estão sendo usados como modelos para estudo de doenças genéticas, sendo que um dos primeiros casos utilizando CRISPR foi desenvolvido este ano. Neste estudo, os autores inseriram uma mutação (A325>T) no gene *SOX10* de porco usando um molde ssODN junto à proteína Cas9 em zigotos isolados de ovidutos (Zhou et al., 2016). A mutação sem sentido p.Arg109Trp nesse gene foi relacionada com o desenvolvimento da síndrome de Waardenburg-Hirschsprung, caracterizada principalmente pela perda de audição e mudanças na coloração do cabelo e pele (Chaoui et al., 2011; Pingault et al., 1998). A transfecção com o molde ssODN alcançou 80% de eficiência, mostrando assim que a alteração foi eficientemente inserida no genoma do porco. Esta mutação também foi detectada nas gônadas dos animais, assemelhando ao que acontece em humanos. Esse modelo passou ser de interesse para o estudo da síndrome de Waardenburg-Hirschsprung e abriu perspectivas para a utilização do porco doméstico no estudo de outras doenças genéticas (Zhou et al., 2016).

Por sua vez, Kang e colegas desenvolveram um modelo em *Macaca fascicularis* (macaco) para a Hipoplasia Adrenal Congênita Associada ao Cromossomo X (AHC), uma doença caracterizada por hipogonadismo. AHC é causada por mutações no gene *DAX1*, importante na formação de órgãos produtores de hormônios. Os pesquisadores desenharam

sgRNAs dirigidos contra o gene *DAX1* e os utilizaram para transfectar, juntamente com Cas9, células de embriões de macacos. Espécimes deficientes de *DAX1* apresentavam alterações no desenvolvimento das glândulas adrenais e testículos acompanhado de uma extensa fibrose, um fenótipo semelhante ao encontrado em humanos com mutações nesse gene, evidenciando que esse modelo pode ser utilizado no estudo da AHC (Kang et al., 2015b; Niu et al., 2014). Atualmente, os modelos em macacos estão despertando interesse, tendo em vista o estudo do efeito de mutações específicas em outros fenótipos patológicos, como o caso do autismo associado a mutações no gene *MeCP2* (Liu et al., 2016; Niu et al., 2014).

O impacto de CRISPR tem sido tão grande que, além dos exemplos já mencionados, outros modelos animais também estão sendo gerados para se estudar doenças genéticas. Como exemplo citamos o furão (*Mustela putorius furo*) para modelar mutações no gene *DCX* associado ao desenvolvimento da lisencefalia (Kou et al., 2015), e as rãs (*Xenopus sp.*) para o estudo de doenças renais (Lienkamp, 2016).

3. Aplicações na medicina

3.1 Combate a patógenos

Identificação e validação de novos alvos terapêuticos

Um dos exemplos dessa aplicação foi conduzido contra o protozoário *Toxoplasma gondii*. Os pesquisadores utilizaram CRISPR para gerar linhagens *knockout* deste microrganismo para o gene da enzima leucina aminopeptidase. Esta proteína pertence à família peptidase M7 e é alvo terapêutico de drogas contra esse patógeno. Quando células Vero (uma linhagem de células de rim do macaco-verde africano) foram infectadas com esta linhagem nocauteada de *T. gondii*, os pesquisadores observaram diminuição na invasão e no desenvolvimento desse patógeno. No modelo *in vivo* em camundongos, eles observaram redução nos efeitos patológicos causados pelo parasita (Zheng et al., 2015). Assim, utilizando CRISPR, os pesquisadores foram capazes validar experimentalmente que a enzima leucina aminopeptidase é importante para o sucesso da infecção do parasita e que, portanto, o desenvolvimento de drogas direcionadas a este alvo são promissoras. É interessante destacar que esse tipo de abordagem pode ser usado tanto para identificar novos alvos terapêuticos quanto para validar potenciais candidatos.

Bloqueio à entrada de vírus

A pesquisa de HIV/AIDS também está sendo incrementada pela CRISPR. Publicações recentes relatam a edição gênica de células do sistema imune humano tanto para impedir a entrada do HIV-1 nos linfócitos T quanto para eliminar o provírus HIV-1 que já estava integrado no genoma dos linfócitos.

Há muito tempo a comunidade científica sabe que indivíduos homozigotos para mutações no correceptor CCR5($\Delta 32$) são refratários a infecção pelo vírus (Liu et al., 1996; Samson et al., 1996), considerando-o, assim, um alvo interessante para controlar a infecção pelo patógeno. Em um dos primeiros trabalhos envolvendo CRISPR e HIV, os pesquisadores selecionaram gRNAs específicos contra este gene e demonstraram que a técnica pode ser

utilizada para gerar linfócitos T CD4⁺ *knockout* para *CCR5*, conferindo a estas células resistência ao HIV-1 (Li et al., 2015a).

Em um trabalho posterior, os autores foram capazes de selecionar células-tronco com pluripotência induzida (iPSCs) homozigotas com os dois alelos *CCR5* na sua forma mutada por CRISPR. Como estas células ainda preservaram sua característica de pluripotência, eles as diferenciaram em macrófagos, que apresentaram resistência a infecção pelo HIV-1, em ensaios de infecção *in vitro* (Kang et al., 2015a).

Um trabalho ainda mais impactante neste contexto foi publicado recentemente por Kaminsk e colaboradores (2016). Os autores utilizaram dois gRNAs diferentes para remover o genoma completo do HIV-1, incluindo todo seu segmento genético entre as sequências 3' e 5' da LTR do provírus integrado em linfócitos T CD4⁺. Os pesquisadores observaram que a coexpressão persistente de Cas9 e dos gRNAs, específica naquelas células que eliminaram o HIV-1, as protegeu contra nova infecção por este vírus. Além disso, CRISPR diminuiu de maneira significativa a replicação e a carga viral em culturas de linfócitos T CD4⁺ de pacientes infectados por HIV-1. Segundo os autores deste trabalho, a edição gênica mediada por CRISPR pode proporcionar uma nova alternativa terapêutica para eliminar o DNA do HIV-1 de linfócitos T CD4⁺ e tem o potencial de servir como uma nova e efetiva plataforma para cura da AIDS (Kaminski et al., 2016).

Outro patógeno de importância para a saúde pública que tem sido objeto de estudo é o vírus da hepatite C (HCV). Lembre-se que a versão mais comum de CRISPR se utiliza da nuclease Cas9 de *Streptococcus pyogenes* (SpCas9), capaz de clivar moléculas de DNA. Contudo, HCV apresenta genoma de ssRNA de polaridade positiva. Neste sentido, os pesquisadores avaliaram se Cas9 derivada da bactéria *Francisella novicida* (FnCas9), que é naturalmente capaz de clivar RNAs bacterianos, seria eficaz no combate ao HCV. Surpreendentemente, os autores obtiveram sucesso com essa estratégia inovadora, demonstrando que CRISPR também é efetiva para o combate de vírus de RNA em células eucarióticas (Price et al., 2015).

3.2 Combate ao câncer

O câncer é uma doença genética somática que leva à multiplicação desregulada de células de um determinado órgão ou tecido, sendo oriunda do acúmulo de mutações no genoma que afetam os proto-oncogenes, modificando-os em oncogenes ativos (White and Khalili, 2016). Segundo a OMS, o número estimado de novos casos de câncer no ano de 2012 no Brasil foi de 14,1 milhões, sendo que 8,2 milhões morreram devido à doença.

Além da identificação do papel de genes em neoplasias, CRISPR também se mostrou como mais uma opção para o tratamento dessa doença, sendo o câncer de mama um caso bem ilustrativo. Neste tipo de câncer, os pesquisadores descreveram os elementos de resposta ao estrogênio distante (DEREs), que englobam cerca de 92 genes. Eles observaram que esses genes apresentavam expressão elevada em pacientes que foram a óbito devido ao câncer de mama (Hsu et al., 2013). Em um trabalho posterior, esses pesquisadores construíram um plasmídeo com dois sgrNAs complementares à região DEREs e o utilizou para transfecção de células de linhagem de câncer de mama MCF-7, resultando na deleção total ou parcial dos genes que compõem este conjunto. Adicionalmente, eles observaram que genes relacionados à proliferação celular foram reprimidos e genes relacionados à apoptose foram induzidos. Portanto, esse trabalho evidenciou o potencial de CRISPR em estratégias para o controle do

crescimento tumoral e, conseqüentemente, aumentar a expectativa de vida dos pacientes (Hsu et al., 2015).

3.3 Terapia genética

O uso de células-tronco com pluripotência induzida (iPSCs) é promissor para terapia de diversas doenças (Takahashi et al., 2007). As iPSCs geradas a partir de fibroblastos de biópsia de pele de pacientes constituem uma alternativa para o tratamento de pessoas que sofrem de Retinite Pigmentosa (RP), um distúrbio hereditário que leva à degeneração progressiva da retina (fotorreceptores), podendo levar à cegueira. Porém, as iPSCs geradas a partir de células de pacientes com RP ainda podem ter alterações genéticas (Schwartz et al., 2015). A utilização de CRISPR para editar os genomas dessas células poderá ajudar a contornar esse problema e permitir a utilização das células dos próprios pacientes no tratamento.

Nesse sentido, Bassuk e colaboradores geraram iPSCs de um paciente com mutação no gene *RPGR* (*retinitis pigmentosa GTPase regulator*), que causa uma variante agressiva de RP ligada ao X (XLRP), visando editá-las via CRISPR para eventual uso futuro em transplante autólogo. Para isso, um gRNA específico para a região mutada no gene *RPGR* foi construído e utilizado para transfecção das iPSCs junto com a endonuclease Cas9 e um DNA doador (ssODN com sequência corrigida do gene *RPGR*). Este DNA atuou como molde para o reparo por recombinação homóloga (HDR), permitindo a edição específica da mutação. Segundo os autores deste trabalho, uma proporção de 13% das células transfectadas apresentaram o reparo esperado na região-alvo, sendo convertidas para o tipo selvagem (Bassuk et al., 2016).

A Distrofia Muscular de Duchenne (DMD) é outro exemplo de doença genética na qual a tecnologia de CRISPR pode ser empregada. A base molecular da DMD é uma mutação no gene da distrofina, localizado no cromossomo X, codificador de uma proteína essencial para o músculo esquelético (Hoffman et al., 1987). Muitas dessas mutações deletérias ocorrem no intervalo entre os éxons 45 a 55. Entretanto, curiosamente, a deleção completa dos éxons 45 a 55 não é deletéria em si, resultando na expressão de uma proteína menor mas funcional (Lu et al., 2011).

Nesse sentido, um grupo de pesquisadores realizou a edição gênica em mioblastos esqueléticos de pacientes com DMD em cultura com o uso do sistema Cas9 e gRNAs para a região específica dos éxons 45-55, que resultou na remoção deste segmento e restauração da matriz de leitura do gene. Isso permitiu a expressão do mRNA e da proteína distrofina truncada que é funcional nessas células (Ousterout et al., 2015). Mais estudos devem ser feitos para garantir a especificidade e excluir a possibilidade de edição do genoma em regiões fora dos alvos.

3.4 Controle de vetores de doenças

CRISPR também permite gerar modificações que prejudiquem o crescimento populacional de vetores de doenças, como o trabalho de Hammond e colegas (2015), que identificaram que alterações nos genes endógenos *AGAP005958*, *AGAP011377* e *AGAP007280* podem levar à esterilidade das fêmeas do mosquito *Anopheles gambiae*. Para conseguir isto, eles desenharam um plasmídeo contendo a Cas9:GFP, um gRNA específico para cada um dos três genes e em cada extremidade do plasmídeo um sítio de recombinação *attB* e o respectivo molde de homologia tanto para a região 5' como para a 3'. O plasmídeo assim construído foi injetado em embriões dos mosquitos. A transmissão desta modificação de

maneira homozigótica teve êxito de até 99%, sendo a alteração no gene *AGAP007280* a que apresentou maior efeito sobre a fertilidade das fêmeas. Esse modelo é muito interessante, tendo em vista a possibilidade de sua utilização para tentar controlar a reprodução desse mosquito na natureza e, conseqüentemente, controlar a transmissão da malária (Hammond et al., 2015).

Outro exemplo muito semelhante que ganhou grande repercussão foi a criação de um mosquito transgênico com resistência ao parasita causador na malária (protozoários do gênero *Plasmodium*) (Gantz et al., 2015), utilizando uma variante de CRISPR denominada MCR (vide cap. 10).

4. Aplicações na agricultura e veterinária

4.1 Combate a pragas agrícolas

Pesquisadores editaram o genoma do trigo (*Triticum aestivum*) com o objetivo de gerar resistência ao oídio do trigo, uma doença causada pelo fungo ectoparasita *Blumeria graminis* f. *sp. tritici*. O alvo para edição foi o gene *MLO*, que codifica uma proteína relacionada à repressão da defesa contra esse fungo em diversas plantas (Büschges et al., 1997). A técnica foi eficiente para gerar o *knockout* de *MLO* e as respectivas plantas com perda da função deste gene demonstraram maior resistência ao fungo (Wang et al., 2014).

Outro estudo com o objetivo de combater pestes agrícolas utilizou CRISPR para provocar mutações no gene *Sxl* em *Drosophila suzukii*, mosca que é praga na fruticultura. O gene *Sxl* é responsável por controlar o desenvolvimento do gênero feminino e pela compensação de dose do cromossomo X. A proteína Cas9 e o gRNA direcionado à região-alvo foram injetados nos embriões desta mosca. Entre os adultos originários deste procedimento, observou-se um fenótipo intersexual em uma frequência de 20%, sendo que esses indivíduos carregavam pelo menos uma mutação no gene *Sxl*. Considerando que as moscas com o fenótipo intersexual não se reproduzem, esse exemplo representa mais uma alternativa de controle da praga, desde que sejam levados em conta os possíveis efeitos à população natural desta espécie de *Drosophila* (Li and Scott, 2015).

4.2 Geração de variação genética – novos cultivares não transgênicos

A mutação é a principal fonte da variabilidade genética e, em consequência, da evolução. Entretanto, o homem interfere nesse processo desde que desenvolvemos a agricultura e a domesticação de animais. Decidimos, segundo nossos interesses, os cruzamentos a serem feitos baseando-nos nas características “desejáveis”, como aumento na produção de grãos, resistência a pestes, ganho de peso em animais de corte, dentre vários outros exemplos. Este processo de melhoria genética por meio de cruzamentos selecionados é feito a partir da variabilidade genética preexistente, *i.e.*, já encontrada na natureza. Agora, por meio de CRISPR, aumentamos imensamente o nosso controle sobre a modificação do DNA. Podemos inserir, deletar ou substituir um ou poucos nucleotídeos de maneira direcionada e controlada, gerando assim novos alelos que são, virtualmente, indistinguíveis da variação genética natural. Em outras palavras, CRISPR permite expandir a variabilidade genética (de plantas, animais), sob a qual nossa seleção artificial pode atuar.

Pyott e colaboradores (2016) utilizaram CRISPR para gerar exemplares de *Arabidopsis thaliana* com mutações pontuais no gene *eIF* (*eukaryotic translation initiation factor*). Sabe-se que plantas homozigotas para mutações neste gene apresentam resistência ao TuMV (*Turnip mosaic virus*), um vírus que causa importantes perdas na agricultura. Assim, após os experimentos de edição genética, os pesquisadores realizaram uma série de cruzamentos controlados para segregar os dois locos (*eIF* e o cassete Cas9/sgRNA). Como resultado, eles obtiveram plantas sem o cassete CRISPR mas mutantes para o gene *eIF*, sendo portanto resistentes ao vírus mas não transgênicas (Pyott et al., 2016). Em especial, estas mutações geradas via CRISPR em *eIF* são indistinguíveis da variação genética natural, aumentando assim as chances de tais plantas serem aceitas pela população.

4.3 Animais resistentes a doenças

Um curioso exemplo da aplicação de CRISPR para se obter animais geneticamente modificados é a edição do genoma de porcos para conferir proteção contra o vírus da síndrome reprodutiva e respiratória dos suínos (PRRSV). Entre as doenças que afetam a suinocultura, esta síndrome é que gera as maiores perdas na América do Norte, Europa e Ásia. Estudos recentes indicam que a proteína CD163 é o receptor que permite a entrada do vírus na célula (Van Breedam et al., 2010). Baseados nisso, Whitworth e colaboradores decidiram utilizar CRISPR para nocautear o gene *CD163* em zigotos fertilizados *in vitro*. Três semanas após o nascimento, os animais derivados deste procedimento foram submetidos à infecção por uma cepa viral de PRRSV e não apresentaram sinais clínicos da doença ou anticorpos contra o vírus durante os 35 dias de monitoramento. De maneira oposta, os animais do grupo controle apresentaram todos os sintomas clássicos e anticorpos durante esse período (Whitworth et al., 2014; Whitworth et al., 2015).

5. Aplicações na indústria

5.1 Produção de biomoléculas

A possibilidade de se alterar o genoma de um organismo para que este sintetize determinados produtos de interesse, tanto para a pesquisa científica como para fins comerciais, já existe há mais de 30 anos. Um exemplo de grande impacto ainda é a insulina humana recombinante produzida pela bactéria *E. coli* (Goeddel et al., 1979). A engenharia genética tem grande interesse na produção de biomoléculas, combustíveis e fármacos. Para isto se faz necessário o conhecimento prévio da modulação das vias metabólicas envolvidas na biossíntese do produto em questão, a fim do melhor aproveitamento dos recursos e aumentar a produção (Li et al., 2015b).

A levedura *Saccharomyces cerevisiae* é uma espécie modelo amplamente utilizada, sendo importante para a produção de alimentos, cerveja, fármacos, produtos de uso industrial e na agricultura e até combustíveis. Ronda e colaboradores (2015) desenvolveram um método de edição gênica chamado CrEdit (*CRISPR/Cas9 mediated genome Editing*), visando integrar várias vias metabólicas em um único cassete de expressão para aumentar a produção de β -caroteno em *S. cerevisiae*.

O sistema consistiu na transfecção de leveduras com Cas9/sgRNAs e três DNAs doadores (um para o gene *crtYB*, outro para *crtI* e outro para *GGPP*). O procedimento levou à clivagem do genoma em três pontos distintos, havendo assim a integração de cada gene em sítios diferentes. Em conjunto, a atividade destas três enzimas resultou na superexpressão de carotenoides na levedura (Ronda et al., 2015).

5.2 Produção de biocombustíveis

A insustentabilidade da utilização a longo prazo de combustíveis de origem fóssil é uma realidade que tem motivado a crescente utilização de biocombustíveis, como o bioetanol. Uma das estratégias que vêm sendo utilizadas para aperfeiçoar a produção de bioetanol é a edição genômica de plantas visando o aumento da síntese do dissacarídeo celobiose. Esse oligossacarídeo é um dos principais compostos utilizados na fabricação de etanol, pois é por meio da sua fermentação que o álcool é produzido nas células vegetais (Galazka et al., 2010).

Os primeiros esforços no sentido de aumentar a produção de celobiose por meio de edição genômica foram realizados por Ryan e colaboradores (2014a), que desenvolveram uma plataforma de edição genômica em *S. cerevisiae* por meio da utilização de plasmídeos contendo a endonuclease Cas9 juntamente com vários sgRNAs. Essa estratégia foi denominada de multiplex CRISPR (mCRISPR) e passou a permitir a indução simultânea de várias quebras no genoma de leveduras seguidas da introdução de fragmentos de DNA interesse nessas quebras.

Posteriormente, Ryan e colaboradores (2014b) utilizaram essa plataforma para modificar a proteína transportadora CDT-1. Esse polipeptídeo é responsável pelo transporte de moléculas de celobiose para dentro das células da levedura, onde elas são convertidas em álcool pelo processo de fermentação. Após apenas uma rodada de edição utilizando a mCRISPR foi possível isolar uma proteína CDT-1 modificada capaz de aumentar em mais de dez vezes o processo de fermentação e consequentemente a produção de álcool. Assim, pode-se afirmar que a utilização de mCRISPR representa uma revolução para a produção de combustíveis produzidos a partir de biomassa e possui potencial para ser utilizada na produção em larga escala do bioetanol.

5.3 Produção de biomateriais

Além de todas as aplicações previamente citadas, a utilização da tecnologia CRISPR possibilita também a manipulação de circuitos biológicos que podem resultar na produção de materiais sintéticos, a partir de biomateriais como a sílica produzida por algas diatomáceas (Hsu et al. 2014). Essa sílica tem o potencial para ser usada na produção de diversos produtos como pomadas dermatológicas, pasta de dente, filtros de água, isolantes térmicos e acústicos, tintas, etc. Dessa forma, o desenvolvimento de estratégias de edição genômica que aumentem a produção de sílica nessas algas permitirá avanços significativos na nanotecnologia aplicada à engenharia de biomateriais (Bozarth et al., 2009; Lebeau and Robert, 2003).

Nesse sentido, Nymark e colaboradores (2016) desenvolveram uma plataforma CRISPR adaptada para produzir *knockout* de genes de forma precisa na alga *Phaeodactylum tricorutum*. A fim de realizar uma prova de conceito da eficiência dessa plataforma, foi construído um vetor expressando a endonuclease Cas9 e um sgRNA contra o gene *CpSRP54*. Como regiões regulatórias de Cas9 e sgRNA, foram utilizadas a sequência promotora do gene *LHCF2* e o promotor U6, respectivamente. Após a edição, os mutantes foram identificados

por meio de ensaios de PCR combinados com *high resolution melting* (HRM), confirmados por sequenciamento e finalmente validados por meio de análises funcionais. A obtenção de indivíduos mutantes após essa triagem confirmou a eficiência dessa plataforma de edição gênica em microalgas, representando um *breakthrough* para a futura utilização da plataforma CRISPR na produção de inúmeros bioprodutos a partir desses microrganismos.

6. Conclusões

Podemos entender as “aplicações de CRISPR” como a “conversão de seus efeitos moleculares em produtos e processos de interesse humano”. Neste capítulo vimos diversos desses exemplos, mas note o leitor que a vasta maioria deles derivou apenas de um dos efeitos moleculares: a edição gênica. Portanto, a real gama de aplicações é imensamente maior ao considerarmos todas as outras possíveis abordagens (marcação de DNA, regulação da expressão gênica, clivagem do RNA, mapeamento de genes, *RNA tracking*, entre outros vindouros).

Referências

- Aida T, Imahashi R, Tanaka K. Translating human genetics into mouse: the impact of ultra-rapid in vivo genome editing. *Dev Growth Differ*. 2014;56:34-45.
- Armstrong GAB, Liao M, You Z, Lissouba A, Chen BE, Drapeau P. Homology Directed Knockin of Point Mutations in the Zebrafish *tardbp* and *fus* Genes in ALS Using the CRISPR/Cas9 System. *PLoS One*. 2016;11:e0150188.
- Assis A, Oliverira E, Donate P, Giulatti S, Nguyen C, Passos G. What is the transcriptome and how it is evaluated? In Passos GA, editor. *Transcriptomics in Health and Disease*. Ribeirão Preto: Springer International Publishing; 2014.
- Bassuk AG, Zheng A, Li Y, Tsang SH, Mahajan VB. Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells. *Sci Rep*. 2016;6:19969.
- Bikard D, Hatoum-Aslan A, Mucida D, Marraffini LA. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe*. 2012;12:177-86.
- Bozarth A, Maier UG, Zauner S. Diatoms in biotechnology: modern tools and applications. *Appl Microbiol Biotechnol*. 2009;82:195-201.
- Büschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, et al. The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell*. 1997;88:695-705.
- Champer J, Buchman A, Akbari OS. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat Rev Genet*. 2016;17:146-59.
- Chaoui A, Watanabe Y, Touraine R, Baral V, Goossens M, Pingault V, et al. Identification and functional analysis of *SOX10* missense mutations in different subtypes of Waardenburg syndrome. *Hum Mutat*. 2011;32:1436-49.
- Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*. 2013;23:1163-71.
- Dow LE. Modeling Disease In Vivo With CRISPR/Cas9. *Trends Mol Med*. 2015;21:609-21.
- Doyle A, McGarry MP, Lee NA, Lee JJ. The construction of transgenic and gene knockout/knockin mouse models of human disease. *Transgenic Res*. 2012;21:327-49.
- Ecsedi M, Rausch M, Großhans H. The let-7 microRNA directs vulval development through a single target. *Dev Cell*. 2015 Feb;32(3):335-44.
- Förster H, McGhee GC, Sundin GW, Adaskaveg JE. Characterization of Streptomycin Resistance in Isolates of *Erwinia amylovora* in California. *Phytopathology*. 2015;105:1302-10.
- Galazka JM, Tian C, Beeson WT, Martinez B, Glass NL, Cate JH. Cellodextrin transport in yeast for improved biofuel production. *Science*. 2010;330:84-6.
- Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc Natl Acad Sci USA*. 2015;201521077.
- Gantz, V.M., and Bier, E. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*. 2015;348:442-4.
- Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, et al. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc Natl Acad Sci USA*. 1979;76:106-10.
- Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat Biotechnol*. 2015;34:78-83.

- Heussler GE, Cady KC, Koeppen K, Bhujus S, Stanton BA, O'Toole GA. Clustered Regularly Interspaced Short Palindromic Repeat-Dependent, Biofilm-Specific Death of *Pseudomonas aeruginosa* Mediated by Increased Expression of Phage-Related Genes. *MBio*. 2015;6:e00129-15.
- Hoffman EP, Brown RH, Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy loco. *Cell*. 1987;51:919-28.
- Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157(6):1262-78.
- Hsu P-Y, Hsu H-K, Hsiao T-H, Ye Z, Wang E, Profit AL, et al. Spatiotemporal control of estrogen-responsive transcription in ER α -positive breast cancer cells. *Oncogene*. 2015
- Hsu P-Y, Hsu H-K, Lan X, Juan L, Yan PS, Labanowska J, et al. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. *Cancer Cell*. 2013;24:197-212.
- Instituto Nacional de Câncer Jose Alencar Gomes da Silva – INCA. Estimativa 2014: Incidência de Câncer no Brasil. Rio de Janeiro; 2014.
- Jia Y, Chen L, Ma Y, Zhang J, Xu N, Liao DJ. To Know How a Gene Works, We Need to Redefine It First but then, More Importantly, to Let the Cell Itself Decide How to Transcribe and Process Its RNAs. *Int J Biol Sci*. 2015;11:1413-23.
- Kaminski R, Chen Y, Fischer T, Tedaldi E, Napoli A, Zhang Y, et al. Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. *Sci Rep*. 2016;6:22555.
- Kang H, Minder P, Park MA, Mesquitta W-T, Torbett BE, Slukvin II. CCR5 Disruption in Induced Pluripotent Stem Cells Using CRISPR/Cas9 Provides Selective Resistance of Immune Cells to CCR5-tropic HIV-1 Virus. *Mol. Ther. Acids*. 2015a;4:e268.
- Kang Y, Zheng B, Shen B, Chen Y, Wang L, Wang J, et al. CRISPR / Cas9-mediated Dax1 knockout in the monkey recapitulates human AHC-HH. 2015b;24:7255-64.
- Kim HD, Shay T, O'Shea EK, Regev A. Transcriptional regulatory circuits: predicting numbers from alphabets. *Science*. 2009;325:429-32.
- Kim J-S, Cho D-H, Park M, Chung W-J, Shin D, Ko KS, et al. CRISPR/Cas9-Mediated Re-Sensitization of Antibiotic-Resistant *Escherichia coli* Harboring Extended-Spectrum β -Lactamases. *J Microbiol Biotechnol*. 2016;26:394-401.
- Kou Z, Wu Q, Kou X, Yin C, Wang H, Zuo Z, et al. CRISPR/Cas9-mediated genome engineering of the ferret. *Cell Res*. 2015;25:1372-5.
- Kraft K, Geuer S, Will AJ, Chan WL, Paliou C, Borschiwer M, et al. Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. *Cell Reports*. 2015;10:833-9.
- Lander ES. The Heroes of CRISPR. *Cell*. 2016;164:18-28.
- Lebeau T, Robert JM. Diatom cultivation and biotechnologically relevant products. Part II: current and putative products. *Appl Microbiol Biotechnol*. 2003;60:624-32.
- LePage DF, Conlon RA. Animal models for disease: knockout, knock-in, and conditional mutant mice. *Methods Mol Med*. 2006;129:41-67.
- Li C, Guan X, Du T, Jin W, Wu B, Liu Y, et al. Inhibition of HIV-1 infection of primary CD4⁺ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J Gen Virol*. 2015a;96:2381-93.
- Li F, Scott MJ. CRISPR/Cas9-mediated mutagenesis of the white and Sex lethal locos in the invasive pest, *Drosophila suzukii*. *Biochem Biophys Res Commun*. 2015;469:911-6.
- Li Y, Lin Z, Huang C, Zhang Y, Wang Z, Tang Y, et al. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. *Metab Eng*. 2015b;1-9.
- Lienkamp SS. Using *Xenopus* to study genetic kidney diseases. *Semin Cell Dev Biol*. 2016;51:117-24.
- Lin S, Ewen-Campen B, Ni X, Housden BE, Perrimon N. In Vivo Transcriptional Activation Using CRISPR/Cas9 in *Drosophila*. *Genetics*. 2015;201:433-42.
- Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*. 1996;86:367-77.
- Liu Z, Li X, Zhang J, Cai Y, Cheng T, Cheng C, et al. Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. *Nature*. 2016;530:98-102.
- Lu Q-L, Yokota T, Takeda S, Garcia L, Muntoni F, Partridge T. The Status of Exon Skipping as a Therapeutic Approach to Duchenne Muscular Dystrophy. *Mol Ther*. 2011;19:9-15.
- Marraffini LA. CRISPR-Cas immunity against phages: its effects on the evolution and survival of bacterial pathogens. *PLoS Pathog*. 2013;9:e1003765.
- Mojica FJM, Juez G, Rodriguez-Valera F. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Mol Microbiol*. 1993;9:613-21.
- Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*. 2014;156:836-43.
- Nymark M, Sharma AK, Sparstad T, Bones AM, Winge P. A CRISPR/Cas9 system adapted for gene editing in marine algae. *Sci Rep*. 2016;6:24951.
- Ostria-Hernández ML, Sánchez-Vallejo CJ, Ibarra JA, Castro-Escarpulli G. Survey of clustered regularly interspaced short palindromic repeats and their associated Cas proteins (CRISPR/Cas) systems in multiple sequenced strains of *Klebsiella pneumoniae*. *BMC Res Notes*. 2015;8:332.
- Ousterout DG, Kabadi AM, Thakore PI, Majoros WH, Reddy TE, Gersbach C. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat Commun*. 2015;6:6244.
- Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, et al. A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell*. 2015;162:675-86.
- Pereira TC. Introdução à técnica de Interferência por RNA – RNAi. Ribeirão Preto: Sociedade Brasileira de Genética; 2013.
- Pingault V, Bondurand N, Kuhlbrodt K, Goerich DE, Préhu MO, Puliti A, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. *Nat Genet*. 1998;18:171-3.

- Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockout mice for genome editing and cancer modeling. *Cell*. 2014;159:440-55.
- Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS. Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc Natl Acad Sci USA*. 2015;112:6164-9.
- Pyott DE, Sheehan E, Molnar A. Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. *Mol Plant Pathol*. 2016.
- Ronda C, Maury J, Jakočiūnas T, Baallal Jacobsen SA, Germann SM, Harrison SJ, et al. CrEdit: CRISPR mediated multi-locos gene integration in *Saccharomyces cerevisiae*. *Microb Cell Fact*. 2015;14:97.
- Ryan OW, Cate JH. Multiplex engineering of industrial yeast genomes using CRISPRm. *Methods Enzymol*. 2014a;546:473-89.
- Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. *eLife*. 2014b;3.
- Salomon DR. A CRISPR Way to Block PERVs—Engineering Organs for Transplantation. *N Engl J Med*. 2016;374:1089-91.
- Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*. 1996;382:722-5.
- Schwartz SD, Regillo CD, Lam BL, Elliott D, Rosenfeld PJ, Gregori NZ, et al. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet*. 2015;385:509-16.
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science*. 2014;343:84-7.
- Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet*. 2015;16:299-311.
- Simoff I, Karlgren M, Backlund M, Lindström A-C, Gaugaz FZ, Matsson P, et al. Complete Knockout of Endogenous Mdr1 (Abcb1) in MDCK Cells by CRISPR-Cas9. *J Pharm Sci*. 2016;105:1017-21.
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007;131:861-72.
- Van Breedam W, Delputte PL, Van Gorp H, Misinzo G, Vanderheijden N, Duan X, et al. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. *J Gen Virol*. 2010;91:1659-67.
- Wang X, Tang Y, Lu J, Shao Y, Qin X, Li Y, et al. Characterization of novel cytochrome P450 2E1 knockout rat model generated by CRISPR/Cas9. *Biochem Pharmacol*. 2016;105:80-90.
- Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol. Advance On*. 2014:1-6.
- White MK, Khalili K. CRISPR/Cas9 and cancer targets: future possibilities and present challenges. *Oncotarget*. 2016.
- Whitworth KM, Lee K, Benne JA, Beaton BP, Spate LD, Murphy SL, et al. Use of the CRISPR/Cas9 System to Produce Genetically Engineered Pigs from In Vitro-Derived Oocytes and Embryos. *Biol Reprod*. 2014;91:78.
- Whitworth KM, Rowland RRR, Ewen CL, Tribble BR, Kerrigan MA, Cino-Ozuna AG, et al. Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. *Nat Biotechnol*. 2015;34:6-8.
- Woo JW, Kim J, Kwon S, Corvalán C, Cho SW, Kim H, et al. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol*. 2015;33:1162-4.
- Xue HY, Ji LJ, Gao AM, Liu P, He JD, Lu XJ. CRISPR-Cas9 for medical genetic screens: applications and future perspectives. *J Med Genet*. 2015;1-7.
- Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*. 2014;514:380-4.
- Yang L, Güell M, Niu D, George H, Leshia E, Grishin D, et al. Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science*. 2015;350:1101-4.
- Yuan L, Sui T, Chen M, Deng J, Huang Y, Zeng J, et al. CRISPR/Cas9-mediated GJA8 knockout in rabbits recapitulates human congenital cataracts. *Sci Rep*. 2016;6:22024.
- Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 2011;29:149-53.
- Zheng J, Jia H, Zheng Y. Knockout of leucine aminopeptidase in *Toxoplasma gondii* using CRISPR/Cas9. *Int J Parasitol*. 2015;45:141-8.
- Zhou X, Wang L, Du Y, Xie F, Li L, Liu Y, et al. Efficient Generation of Gene-Modified Pigs Harboring Precise Orthologous Human Mutation via CRISPR/Cas9-Induced Homology-Directed Repair in Zygotes. *Hum Mutat*. 2016;37:110-8.

Comparação com outras técnicas para *knockout* e *knockdown*: RNAi, *Gene Targeting*, ZFNs e TALENs

Capítulo

5

Leonardo Chicaybam^{1,2}, Bárbara Peixoto¹ e Martin Bonamino^{1,2}

¹*Programa de Carcinogênese Molecular, Instituto Nacional do Câncer, Rio de Janeiro, RJ - Brasil.*

²*Vice-presidência de Pesquisa e Laboratórios de Referência, Fundação Instituto Oswaldo Cruz, Rio de Janeiro, RJ - Brasil.*

Estruturação do capítulo

1. Introdução
2. Interferência por RNA (RNAi)
3. *Gene Targeting* e Cre/loxP
4. Nucleases Dedo de Zinco (ZFNs)
5. *Transcription activator-like effector nucleases* (TALENs)
6. Conclusões

1. Introdução

Antes do surgimento da técnica de CRISPR foram desenvolvidas ferramentas como RNAs antissenso, ribozimas, RNAs de interferência, nucleases dedo de zinco (*zinc finger nucleases*, ZFNs) e TALENs, além dos métodos clássicos de geração de células e animais nocaute (*e.g.*, *Gene Targeting*), que permitiram a manipulação genética de diversos organismos, viabilizando os primeiros experimentos de supressão e edição de genes. Focaremos a seguir nas ferramentas que se consolidaram para a dissecação genética, iniciando com as técnicas que envolvem a modificação direta do transcriptoma da célula, como a Interferência por RNA (RNAi), prosseguindo com as estratégias de alteração do genoma (*Gene Targeting*, ZFNs e TALENs).

2. Interferência por RNA (RNAi)

A **Interferência por RNA (RNAi)** foi a primeira destas ferramentas desenvolvidas que se consolidou amplamente para uso experimental, e os primeiros indícios deste mecanismo foram observados em plantas, onde se constatou a supressão de genes homólogos após a superexpressão do gene *CHS* (*chalcone synthase*) (Napoli et al., 1990). No entanto, o mecanismo por trás deste efeito permaneceu indeterminado e foi apenas descrito oito anos depois por Andrew Fire e Craig Mello (Fire et al., 1998). Utilizando o nematódeo *C. elegans* como modelo, esses pesquisadores demonstraram que moléculas longas de dupla fita de RNA (dsRNA) contendo parte do gene-alvo eram capazes de inibir de forma específica a expressão deste gene através da diminuição do mRNA-alvo. Esse mesmo efeito não aconteceu quando apenas as fitas senso ou antissenso eram injetadas, demonstrando que o mecanismo é dependente da estrutura dupla fita do RNA utilizado.

Trabalhos posteriores elucidaram a fundo o mecanismo de ação e concluíram que ele é altamente conservado desde plantas até mamíferos, embora com pequenas diferenças, envolvendo uma maquinaria de processamento de dsRNA em que diversas proteínas estão envolvidas. Uma dessas proteínas se chama Dicer, um componente crucial da via que reconhece e processa o dsRNA, clivando a longa molécula em pequenos duplexes de 21 nucleotídeos contendo uma ponta não pareada de dois nucleotídeos na porção 3', também chamados de siRNAs (*small interfering RNAs*). Subsequentemente, uma proteína chamada Argonata se junta ao complexo e se liga a uma das fitas do siRNA, enquanto a outra será removida. Após a associação de componentes adicionais, este complexo passa a se chamar RISC (*RNA-induced silencing complex*), sendo capaz de interagir com o mRNA-alvo através do pareamento com a fita derivada do siRNA (Dillon et al., 2005). No caso de um pareamento completo entre o RISC e o mRNA-alvo (mecanismo mais comum em plantas), o mRNA é clivado pelo RISC e a expressão é abolida. Já o pareamento incompleto induz um bloqueio da maquinaria de tradução, uma vez que o RISC permanece ligado ao mRNA-alvo, sendo este mecanismo mais comum nos vertebrados (figura 1). Trabalhos posteriores mostraram que transfecção de siRNAs (em torno de 21 nucleotídeos) é a mais indicada por induzirem a supressão do mRNA-alvo sem ativar sensores celulares de RNAs dupla fita longos como a proteína PKR, possibilitando a aplicação desta técnica em mamíferos (Elbashir et al., 2001).

A RNAi se demonstrou uma técnica muito útil, pois induz a supressão da expressão gênica de forma rápida, fácil e com especificidade, facilitando o estudo de genes que ainda não têm a função descrita. Por ser uma molécula pequena, a síntese do siRNA é barata e os

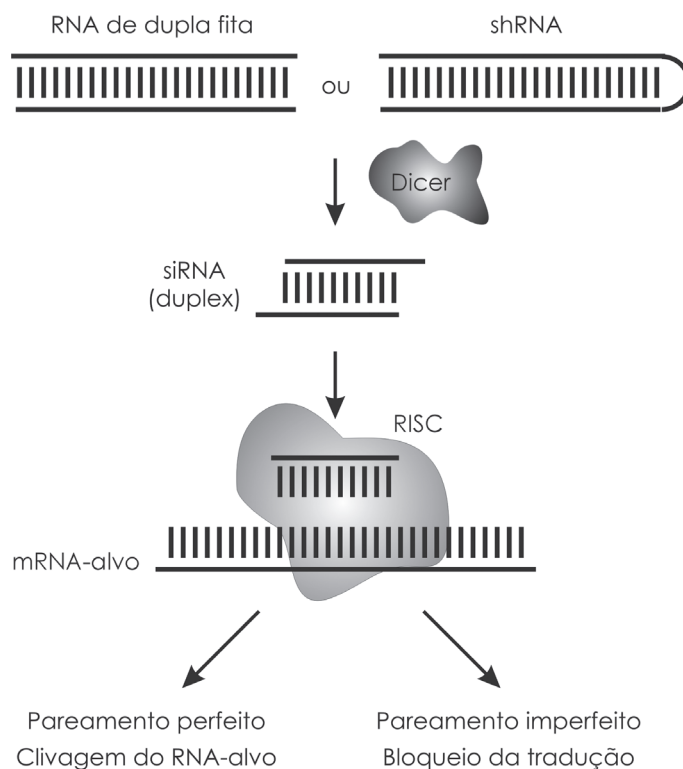


Figura 1. Via de processamento da RNAi. Os RNAs de dupla fita (dsRNAs) ou na forma de grampos (shRNAs) são clivados pela Dicer dando origem ao siRNA, um RNA dupla fita de 21 nucleotídeos contendo dois nucleotídeos não pareados na porção 3'. Uma das fitas do siRNA é transferida para RISC, guiando-o até reconhecer o mRNA-alvo através de complementaridade. O pareamento perfeito leva à degradação do alvo, enquanto o pareamento imperfeito bloqueia sua tradução. Imagem: Pereira, TC (2016).

métodos de transfecção (eletroporação, lipofecção) são muito eficientes, ampliando o número de tipos celulares e organismos em que essa técnica pode ser utilizada. Além disso, diversos pesquisadores exploraram o potencial terapêutico dessas moléculas, sendo o exemplo mais avançado o uso da RNAi para o tratamento de amiloidose familiar por transtirretina (Coelho et al., 2013). A utilização dessa técnica resultou em avanços científicos em diversas áreas, tendo sua importância reconhecida com a outorga do Prêmio Nobel de Medicina em 2006 a Fire e Mello.

No entanto, o uso da técnica de RNAi também possui algumas desvantagens que foram sendo descobertas ao longo dos anos. Observou-se que o silenciamento gênico frequentemente não é completo (Krueger et al., 2007), gerando assim um efeito de *knockdown* (e não de *knockout*). Adicionalmente, a maquinaria de processamento parece estar mais ativa no citoplasma, *i.e.*, o silenciamento de transcritos nucleares é menos eficiente (Derrien et al., 2012). Efeitos inespecíficos (ou seja, que não sobre o mRNA-alvo) também foram descritos e são mais frequentes em situações em que a complementaridade entre o RISC e o mRNA não é completa (Jackson et al., 2003; Jackson et al., 2006). Foi demonstrado também que esses efeitos inespecíficos parecem ser dependentes da dosagem utilizada (Wang et al., 2009), e

um estudo demonstrou que os fenótipos derivados desses efeitos pode prevalecer sobre o efeito no mRNA-alvo (Franceschini et al., 2014). Outro fato importante é que o uso da RNAi em altas doses pode sobrecarregar a via de processamento, prejudicando a biogênese de microRNAs endógenos (que compartilham a mesma via) e consequentemente alterando os níveis de diferentes transcritos e o fenótipo da célula (Khan et al., 2009). Vale lembrar que a supressão mediada por RNAi acontece enquanto houver moléculas de siRNA disponíveis para serem processadas, sendo assim o silenciamento é transitório (geralmente alguns poucos dias). Para abordagens em que é necessário o silenciamento por longos períodos foram desenvolvidos RNAs com estrutura em grampos (*short hairpins* RNAs – shRNAs), que são processados pela mesma via da RNAi. Os shRNAs podem ser inseridos em vetores retrovirais ou outros vetores de expressão permanente, permitindo a integração do conjunto promotor:shRNA no genoma e consequentemente a expressão estável do shRNA, levando à supressão constante do mRNA-alvo.

Em síntese, a RNAi é uma ferramenta que permite o silenciamento gênico pós-transcricional, temporário, manipulando assim a abundância do transcrito alvo sem afetar a sequência nucleotídica do gene correspondente. Essa redução parcial dos níveis de RNA é conhecida como *knockdown*.

3. *Gene Targeting* e Cre/loxP

A utilização de células e animais nocaute (com a anulação do gene de interesse) sempre teve um papel importante no avanço do entendimento de sistemas biológicos, propiciando o estudo da função de diversos genes. A tecnologia que viabilizou a indução desses nocautes é denominada ***Gene Targeting*** e foi descrita na década de 1980, quando um estudo demonstrou a possibilidade de induzir recombinação homóloga entre um DNA exógeno e sequências homólogas existentes no genoma de células de mamíferos (Smithies et al., 1985). Para que a recombinação ocorra é necessário a construção de um vetor plasmidial contendo um marcador para seleção, frequentemente o gene que confere resistência ao antibiótico G418 (gene *Neo^R*), flanqueado por regiões de homologia com o sítio-alvo para a inserção. Após a transfecção deste vetor na célula de interesse, as regiões de homologia induzem a recombinação com o sítio homólogo no genoma, inserindo o gene *Neo^R* no local designado. Como essas regiões de homologia no vetor são desenhadas de forma que o *Neo^R* seja inserido dentro do gene-alvo, a expressão deste último é abolida. Esse procedimento também pode ser realizado com sucesso em células-tronco embrionárias de camundongo e, após injeção em blastocistos, essas células podem dar origem a diversos tecidos, dentre eles a linhagem germinativa. Após cruzamentos dos animais contendo as alterações de interesse, uma linhagem pura é gerada, na qual os animais possuem a deleção do gene em todas as células.

Apesar de muito útil, esse método é pouco eficiente. Estima-se que a recombinação ocorra corretamente em apenas uma célula em cada 10^5 - 10^7 (Sargent et al., 1998). Além disso, a alta frequência de integrações randômicas (não homólogas) dificulta a geração de animais com a alteração desejada, fazendo com que muitas células ou animais tenham que ser analisados (Roth et al., 1988). Foi descrito que a utilização de uma região de homologia maior (Thomas et al., 1987) e o uso de métodos de enriquecimento com seleção positiva e negativa (Mansour et al., 1988) podem aumentar muito a eficiência da técnica, contornando esses fatores. Mesmo com a baixa eficiência, esta técnica foi utilizada para a geração de

diversos animais nocauteados, contribuindo para a elucidação de diversos mecanismos celulares (Galli-Taliadoros et al., 1995).

No entanto, se o gene de interesse tiver um papel importante no desenvolvimento do animal, a alteração pode resultar em letalidade embrionária. Além disso, em determinados casos pode ser importante que a alteração no gene esteja presente apenas em determinados tipos celulares ou em certa fase da vida do animal. Para esses casos foi desenvolvido o sistema de recombinação **Cre/loxP**, baseado na recombinase *Cre* derivada do bacteriófago P1 (Gu et al., 1994). Nesse sistema é necessário que duas linhagens de animais sejam geradas previamente: um camundongo expressando a recombinase *Cre* no tipo celular de interesse, através do uso de um promotor tecido-específico; e um camundongo possuindo o gene a ser alterado flanqueado por sítios *loxP*. Os sítios *loxP* são assimétricos e a orientação destes define o resultado da recombinação, podendo induzir a inversão, translocação ou deleção do gene. A *Cre* reconhece os sítios *loxP*, catalisando a reação de recombinação. Após cruzamento das duas linhagens, camundongos contendo a alteração apenas no tecido onde *Cre* é expressa são gerados. Um refinamento da técnica usando sistemas induzíveis de expressão gênica permitiram um maior controle temporal da expressão de *Cre*, aumentando a flexibilidade e possíveis aplicações do sistema através da indução da expressão de *Cre* por drogas fornecidas aos animais, como o tamoxifeno (Kunh et al., 1995; Feil et al., 1996). A *Cre* recombinase pode ainda ser dada de forma exógena às células, através de métodos de entrega como a utilização de adenovírus recombinantes carreando o cDNA desta enzima (Kanegae et al., 1995).

Métodos que permitam a realização de cortes em sequências específicas de DNA têm o potencial de *aumentar em muito as taxas de recombinação homóloga usando um DNA doador*. Entre as técnicas que permitem essa edição de DNA sítio dirigida estão as ferramentas descritas a seguir: ZFNs, TALENs e CRISPR.

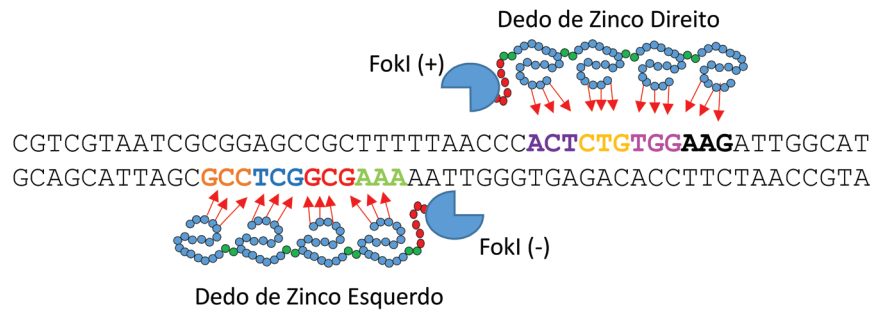
4. Nucleases Dedo de Zinco (ZFNs)

Paralelamente ao desenvolvimento da RNAi, pesquisadores começaram a desenvolver ferramentas para edição genômica, permitindo a manipulação e alteração da sequência de nucleotídeos de forma específica através do uso de nucleases que atuem no núcleo da célula. As nucleases são capazes de promover uma quebra na dupla fita de DNA, induzindo as vias de reparo por NHEJ ou HDR e permitindo, respectivamente, a inserção/deleção de bases e a recombinação homóloga com o DNA doador. O grande desafio desta abordagem é o direcionamento preciso da nuclease para o sítio de interesse, evitando quebras no DNA em outras regiões. Desde a década de 1970, enzimas capazes de reconhecer e clivar determinadas sequências de DNA, como as enzimas de restrição, já estavam disponíveis, mas as sequências de reconhecimento eram compostas de, no máximo, oito nucleotídeos apenas, tornando a manipulação de grandes segmentos de DNA um desafio devido à alta frequência de sítios-alvo em genomas. Estudos demonstraram que em alguns casos, como na nuclease FokI, o reconhecimento do DNA e o sítio catalítico estão separados em diferentes domínios (Li et al., 1992). Este trabalho foi muito importante, pois permitiu a criação de proteínas quiméricas contendo domínios de ligação ao DNA e o sítio catalítico da FokI, propiciando a clivagem de uma região do genoma com alta especificidade. O domínio escolhido para ser utilizado foi o dedo de zinco (*zinc finger*), que possui esse nome devido à sua estrutura de alfa hélice e à necessidade de um átomo de zinco para essa estrutura se formar (Kim et al., 1996). Esses

domínios são modulares e cada um deles é capaz de reconhecer uma trinca de nucleotídeos de forma específica (porém com diferentes níveis de seletividade), possibilitando a criação de uma cadeia polipeptídica contendo vários domínios dedos de zinco em sequência que guiam a nuclease para o sítio-alvo (figura 2a). Como a homodimerização de FokI é necessária para que ocorra a clivagem (Bitinaite et al., 1998), duas proteínas quiméricas devem se ligar ao sítio-alvo com um espaçamento definido entre elas (5-7 pb) para que a manipulação genética seja eficiente. Este requisito faz com que a especificidade do sistema usando *nucleases dedo de zinco* (ZFNs) seja maior do que a especificidade de ferramentas que usam apenas um sítio de reconhecimento (CRISPR, quando utilizando apenas uma Cas9), uma vez que os dois monômeros precisam interagir corretamente com o sítio-alvo para que ocorra a clivagem (Kim et al., 2014).

As ZFNs foram as primeiras ferramentas capazes de induzir uma alteração de forma controlada em um local predeterminado no genoma e já foram utilizadas com sucesso para inserção do gene *IL2RG* em modelos de imunodeficiência ligada ao X (SCID-X1) *in vitro* (Urnov et al., 2005) e *in vivo* (Genovese et al., 2014), e também na indução de resistência à infecção por HIV através da deleção simultânea dos genes *CCR5* e *CXCR4* (Didigu et al., 2014). Esta

(A) Nucleases Dedo de Zinco



(B) TALENs

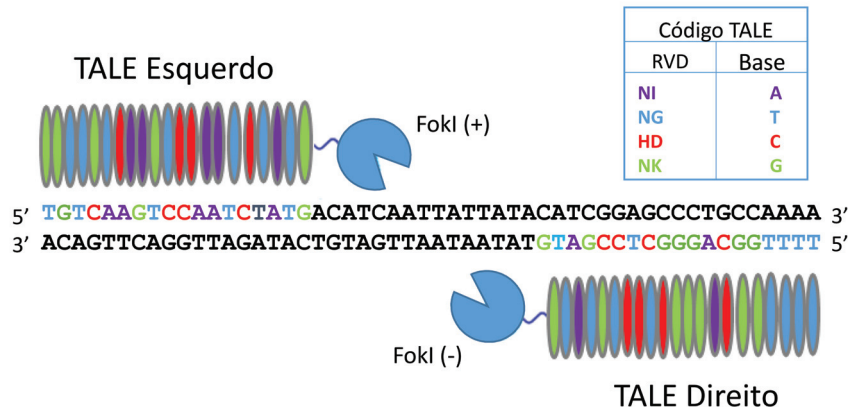


Figura 2. Esquema representativo do reconhecimento do DNA-alvo pelas nucleases. (A) Dedo de zinco e (B) TALENs. Assim como CRISPR, as duas tecnologias induzem a quebra da dupla fita de DNA, recrutando a maquinaria de reparo ao DNA e permitindo a alteração genética. Imagem: Chicaybam, L; Peixoto, B e Bonamino, M (2016).

última abordagem já vem sendo explorada clinicamente em pacientes com HIV, nos quais a edição do gene *CCR5* em linfócitos T CD4+ se mostrou segura e as células editadas (*CCR5* negativas) apresentaram menor declínio no sangue quando comparada às não editadas após a suspensão transitória do tratamento antirretroviral (Tebas et al., 2014).

Apesar dos resultados promissores, a tecnologia das ZFNs demanda uma complexa engenharia no desenho e na montagem dos domínios dedos de zinco, além de problemas de especificidade. Não existem domínios descritos que reconheçam as 64 combinações possíveis de trincas de nucleotídeos, limitando a escolha do sítio-alvo (Segal et al., 1999). Nem todas as ZFNs montadas são capazes de clivar o DNA de forma eficiente, e aparentemente sítios-alvo contendo guanina na região 5' apresentam maior taxa de sucesso (Kim et al., 2009). Alguns estudos mostraram que esses domínios podem ser montados a partir de dedos de zinco pré-caracterizados (ou seja, com interação confirmada com determinada trinca), mas frequentemente quando colocados na construção final não possuem atividade (Ramirez et al., 2008) ou apresentam citotoxicidade devido à clivagem inespecífica (Cornu et al., 2008). Um estudo demonstrou que a clivagem inespecífica pode acontecer quando apenas um monômero se liga ao sítio-alvo. A utilização de uma FokI modificada para formar obrigatoriamente heterodímeros diminui este efeito, diminuindo também a citotoxicidade (Miller et al., 2007; Doyon et al., 2011). Estas características tornam o processo de desenho e validação das ZFNs demorado e custoso.

5. *Transcription activator-like effector nucleases (TALENs)*

A tecnologia do **TALEN** foi descrita em 2009 (Boch et al., 2009; Moscou and Bogdanove, 2009) e possui bastante similaridade com as ZFNs, usando o mesmo domínio FokI para a clivagem do sítio-alvo. Assim como as outras ferramentas citadas, o TALEN é utilizado para manipulação genética de diversos organismos e espécies, desde vírus (Bloom et al., 2013) e leveduras (Christian et al., 2010; Li et al., 2011) a células de mamíferos (Gaj, 2014; Segal and Meckler, 2013). A sua estrutura de ligação ao DNA é composta por domínios conhecidos como *transcription activator-like effectors* (TALEs), que são provenientes de um patógeno que acomete plantas, a bactéria *Xanthomonas* spp (Li et al., 2011).

Os TALEs são formados por arranjos em sequências de 33-35 resíduos de aminoácidos que, por sua vez, reconhecerão um único par de bases no sulco maior (figura 2b) (Mak et al., 2012). Nas posições 12 e 13 deste arranjo estão presentes os dois aminoácidos, chamados Dirresíduos Variáveis Repetidos (*repeat variable diresidues* – RVDs), que determinam a especificidade do reconhecimento do nucleotídeo na ligação ao DNA (Boch et al., 2009). Para o desenho dessa ferramenta de edição, alguns programas podem ser utilizados, como o E-TALEN (Heigwer et al., 2013) ou TALE-NT 2.0 (Doyle et al., 2012). Após os ZFNs, o TALEN surgiu como uma ferramenta mais específica, estável, eficiente, menos tóxica e de desenho mais fácil (Guilinger et al., 2014; Hockemeyer et al., 2011; Mussolino et al., 2011). Não se sabe ao certo o motivo de os TALENs apresentarem maior especificidade (Mussolino and Cathomen, 2012).

Os parâmetros que definem a atividade *off-target* da nuclease, tanto para o TALEN quanto para o ZFN, são: (i) a especificidade de ligação ao seu alvo (Pattanayak et al., 2011; Richard Gabriel et al., 2011), (ii) o potencial intrínseco de dimerização do homodímero da nuclease FokI (Richard Gabriel et al., 2011) e (iii) o promíscuo reconhecimento de outras

regiões que não a região-alvo (Pattanayak et al., 2011; Richard Gabriel et al., 2011). Embora alguns estudos proponham que os ZFNs são mais específicos, sabe-se que ao aumentar o tamanho do TALE há o acréscimo de especificidade (Guo et al., 2010; Yuka Shimizu et al., 2011.). Em paralelo, aumentar o tamanho dos dedos de zinco gera um impacto negativo em sua atividade (Guo et al., 2010; Yuka Shimizu et al., 2011).

No entanto, a construção de um arranjo específico do TALEN é trabalhosa e exige tempo, uma vez que consiste, frequentemente, de aproximadamente 20 RVDs (Holkers et al., 2013). Além dessas ferramentas, outras também foram desenvolvidas com o intuito de gerar desenhos personalizados, tanto para a CRISPR quanto para o TALEN, visando melhorá-las, tornando-as cada vez mais específicas e estáveis (Cermak et al., 2011; Reyon et al., 2013; Xie et al., 2014). A maior limitação ao se desenhar um TALEN é a necessidade da presença de uma timina à extremidade 5' da sequência-alvo para que haja o reconhecimento e posterior clivagem (Mak et al., 2012), limitação que não existe em CRISPR-Cas9. No entanto, recentemente, a geração de variantes do TALEN permite maior variedade de possíveis sítios-alvo, através do reconhecimento de outras bases localizadas à extremidade 5' (Doyle et al., 2013; Lamb et al., 2013). A eficiência de edição do TALEN é de 1 a ~60%, comparável à de CRISPR (2,3 a 79%) (Cho et al., 2013; Cong et al., 2013; Kim et al., 2013; Mali and Norville, 2013; Reyon et al., 2013).

A vantagem da utilização de CRISPR consiste em alguns fatores. Primeiramente, o benefício da simplicidade e facilidade do desenho do gRNA (são necessários apenas cerca de 20 nucleotídeos que irão compor o RNA-guia e serão complementares à região-alvo) (Cong et al., 2013). Além da maior eficiência de mutação, CRISPR também permite que haja a manipulação de mais de um gene ao mesmo tempo, podendo gerar mutações múltiplas (Li et al., 2013; Niu et al., 2014; Wang et al., 2013), sem a problemática dos efeitos *off-target* vista em TALENs e ZFNs (Tabela 1) (Doyle et al., 2012; Öllü et al., 2010).

Tabela 1. Quadro comparativo entre ZFNs, TALENs e CRISPR.

Características	ZFNs	TALENs	CRISPR
Determinantes da especificidade ao DNA	Domínios dedo de zinco	TALE	gRNA
Nuclease	FokI	FokI	Cas9
Taxa média de mutação	Baixa	Alta	Alta
Tamanho do sítio-alvo	18-36 pb	30-40 pb	17-22 pb
Restrições no sítio-alvo	Deve ser rico em G	Começar com T e terminar A	Terminar com NGG (PAM)*
Densidade de sítios-alvo	1 a cada 100 pb	1 a cada 35 pb	1 a cada 8 pb
Efeitos <i>off-target</i>	Alta	Baixa	Variável
Citotoxicidade	Variável para alta	Baixa	Baixa
Tamanho da construção	~1 kb x 2	~3 kb x 2	4,2 kb (Cas9) + 0,1 kb (gRNA)
Possibilidade de construir bibliotecas para edição em escala genômica	Não	Sim, mas tecnicamente difícil	Sim

*PAM de *S. pyogenes*.

6. Conclusões

A edição genética não é uma abordagem nova, tendo suas raízes estabelecidas há mais de três décadas, com os trabalhos de Oliver Smithies e Mario Capecchi envolvendo a integração de um DNA doador via recombinação homóloga (Smithies et al., 1985), que posteriormente receberam o Prêmio Nobel. O uso de nucleases que permitem a clivagem do sítio-alvo (*e.g.*, ZFN, TALENs e CRISPR), gerando quebras de dupla fita e acionando o sistema de reparo, aumenta drasticamente a eficiência deste processo. O grande sucesso de CRISPR se deve ao seu *design* simples, baixo custo e eficiência superior. Considerando que os criadores da RNAi também foram laureados com o Nobel, as chances de os pioneiros de CRISPR serem premiados são elevadíssimas.

Referências

- Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci USA*. 1998;95:10570-5.
- Bloom K, Ely A, Mussolino C, Cathomen T, Arbutnot P. Inactivation of hepatitis B virus replication in cultured cells and in vivo with engineered transcription activator-like effector nucleases. *Mol Ther*. 2013;21:1889-97.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*. 2009;326:1509-12.
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011;39:e82.
- Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*. 2013;31:230-2.
- Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. 2010;186:757-61.
- Coelho T, Adams D, Silva A, Lozeron P, Hawkins PN, Mant T, et al. Safety and efficacy of RNAi therapy for transthyretin amyloidosis. *N Engl J Med*. 2013;369:819-29.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819-23.
- Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichinger M, Joung JK, et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther*. 2008;16:352-8.
- Derrien B, Baumberger N, Schepetilnikov M, Viotti C, De Cillia J, Ziegler-Graff V, et al. Degradation of the antiviral component ARGONAUTE1 by the autophagy pathway. *Proc Natl Acad Sci USA*. 2012;109:15942-6.
- Didigu CA, Wilen CB, Wang J, Duong J, Secreto AJ, Danet-Desnoyers GA, et al. Simultaneous zinc-finger nuclease editing of the HIV coreceptors ccr5 and cxcr4 protects CD4+ T cells from HIV-1 infection. *Blood*. 2014;123:61-9.
- Dillon CP, Sandy P, Nencioni A, Kissler S, Robinson DA, Van Parijs L. RNAi as an experimental and therapeutic tool to study and regulate physiological and disease processes. *Annu Rev Physiol*. 2005;67:147-73.
- Doyle EL, Boohar NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, et al. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res*. 2012;40:W117-22.
- Doyle EL, Hummel AW, Demorest ZL, Starker CG, Voytas DF, Bradley P, et al. TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat-1. *PLoS One*. 2013;8:e82120.
- Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*. 2011;8:74-9.
- Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*. 2001;411:494-8.
- Feil R, Brocard J, Mascrez B, LeMeur M, Metzger D, Chambon P. Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci USA*. 1996;93:10887-90.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998;391:806-11.
- Franceschini A, Meier R, Casanova A, Kreibich S, Daga N, Andrichschke D, et al. Specific inhibition of diverse pathogens in human cells by synthetic microRNA-like oligonucleotides inferred from RNAi screens. *Proc Natl Acad Sci USA*. 2014;111:4548-53.
- Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol*. 2011;29:816-23.
- Gaj T, Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013;31:397-405.
- Galli-Taliadoros LA, Sedgwick JD, Wood SA, Körner H. Gene knock-out technology: a methodological overview for the interested novice. *J Immunol Methods*. 1995;181:1-15.
- Genovese P, Schirotti G, Escobar G, Di Tomaso T, Firrito C, Calabria A, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*. 2014;510:235-40.

- Gu H, Marth JD, Orban PC, Mossmann H, Rajewsky K. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science*. 1994;265:103-6.
- Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014;32:577-82.
- Guo J, Gaj T, Barbas CF 3rd. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol*. 2010;400:96-107.
- Heigwer F, Kerr G, Walther N, Glaeser K, Pelz O, Breinig M, et al. E-TALEN: a web tool to design TALENs for genome engineering. *Nucleic Acids Res*. 2013;41:e190.
- Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*. 2011;29:731-4.
- Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C, et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res*. 2013;41:e63.
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol*. 2003;21:635-7.
- Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L, et al. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*. 2006;12:1179-87.
- Kanegae Y, Lee G, Sato Y, Tanaka M, Nakai M, Sakaki T, et al. Efficient gene activation in mammalian cells by using recombinant adenovirus expressing site-specific Cre recombinase. *Nucleic Acids Res*. 1995;23:3816-21.
- Khan AA, Betel D, Miller ML, Sander C, Leslie CS, Marks DS. Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat Biotechnol*. 2009;27:549-55.
- Kim H, Kim J-S. A guide to genome engineering with programmable nucleases. *Nat Rev Genet*. 2014.
- Kim HJ, Lee HJ, Kim H, Cho SW, Kim J-S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res*. 2009;19:1279-88.
- Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol*. 2013;31:251-8.
- Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci USA*. 1996;93:1156-60.
- Krueger U, Bergauer T, Kaufmann B, Wolter I, Pilk S, Heider-Fabian M, et al. Insights into effective RNAi gained from large-scale siRNA validation screening. *Oligonucleotides*. 2007;17:237-50.
- Kühn R, Schwenk F, Aguet M, Rajewsky K. Inducible gene targeting in mice. *Science*. 1995;269:1427-9.
- Lamb BM, Mercer AC, Barbas CF 3rd. Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucleic Acids Res*. 2013;41:9779-85.
- Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci USA*. 1992;89:4275-9.
- Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res*. 2011;39:359-72.
- Li W, Teng F, Li T, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol*. 2013;31:684-6.
- Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*. 2012;335:716-9.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823-6.
- Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature*. 1988;336:348-52.
- Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. 2007;25:778-85.
- Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009;326:1501.
- Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol*. 2012;23:644-50.
- Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res*. 2011;39:9283-93.
- Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*. 1990;2:279-89.
- Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*. 2014;156:836-43.
- Oliver Smithies RGG, Boggs SS, Koralewski MA. Insertion of DNA sequences into the human chromosomal b-globin locus by homologous recombination. *Nature*. 1985;317:19.
- Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods*. 2011;8:765-70.
- Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods*. 2008;5:374-5.
- Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012;30:460-5.
- Roth D, Wilson J. Illegitimate recombination in mammalian cells. *Genetic Recombination*. 1988;621-653.
- Sargent RG, Wilson JH. Recombination and gene targeting in mammalian cells. *Current Research Molecular Therapy*. 1998;1:584.
- Segal DJ, Dreier B, Beerli RR, Barbas CF. Toward controlling gene expression at will: Selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci USA*. 1999;96:2758-63.

- Segal DJ, Meckler JF. Genome engineering at the dawn of the golden age. *Annu Rev Genomics Hum Genet.* 2013;14:135-58.
- Shimizu Y, Şöllü C, Meckler JF, Adriaenssens A, Zykovich A, Cathomen T, et al. Adding fingers to an engineered zinc finger nuclease can reduce activity. *Biochemistry.* 2011;50:5033-41.
- Şöllü C, Pars K, Cornu TI. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids.* 2010.
- Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *N Engl J Med.* 2014;370:901-10.
- Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell.* 1987;51:503-12.
- Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature.* 2005;435:646-51.
- Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* 2013;153:910-8.
- Wang X, Wang X, Varma RK, Beauchamp L, Magdaleno S, Sendera TJ. Selection of hyperfunctional siRNAs with improved potency and specificity. *Nucleic Acids Res.* 2009;37:e152.
- Xie S, Shen B, Zhang C, Huang X, Zhang Y. sgRNAs9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLOS One.* 2014;9:e100448.

Variantes naturais e artificiais de Cas9 e sgRNAs

Capítulo

6

Julio Cesar Cetrulo Lorenzi¹

¹*The Rockefeller University, NY - EUA*

Estruturação do capítulo

1. Introdução
2. Cas9 clássica de *Streptococcus pyogenes* (SpCas9)
3. *Nickases*
4. Cas9 com sítios catalíticos inativos (*dead* Cas9 - dCas9)
5. Entendimento e desenvolvimento da região PAM
6. Cas9 de alta precisão
7. Sistemas de ativação indireta da Cas9
8. Modificações nos sgRNAs
9. Cpf1
10. Conclusões

1. Introdução

Neste capítulo temos como objetivo relacionar as variantes naturais e artificiais da nuclease Cas9 e do sgRNA, além relacioná-las a alguns aperfeiçoamentos da técnica de CRISPR e aplicações.

2. Cas9 clássica de *Streptococcus pyogenes* (SpCas9)

A primeira demonstração e caracterização de uma endonuclease de DNA guiada por RNAs aconteceu em 2012 (Jinek et al., 2012). Nesse artigo, que se tornou clássico após apenas quatro anos de sua publicação, foi demonstrado que em sistemas CRISPR do Tipo II as proteínas Cas9, nesse caso isoladas de *Streptococcus pyogenes* (SpCas9), dependem de duas espécies de RNA (crRNA e tracrRNA) para realizar sua função como endonuclease (figura 1a). Esse estudo foi fundamental na disseminação do sistema CRISPR-Cas9 uma vez que mostrou detalhadamente os passos necessários para a aplicação deste de maneira simples. Devemos enfatizar também que esse estudo iniciou o aprimoramento dos métodos necessários para o uso da Cas9, como ferramenta na biologia molecular. O primeiro avanço foi a demonstração da possibilidade do uso de uma única molécula de RNA como guia para o funcionamento do sistema (figura 1b). Foi determinado que a molécula RNA-guia (crRNA) pode ser conectada à molécula de RNA ativadora (tracrRNA) por uma cadeia de 4 nucleotídeos (GAAA *tetraloop*), criando uma única molécula de RNA quimérico capaz de ativar o sistema (sgRNA, ou simplesmente gRNA). Essa estratégia simplificou o uso do sistema, uma vez que esse RNA quimérico pode ser codificado e clonado em um único plasmídeo, facilitando a transfecção bacteriana. Sim, nessa época, em 2012, ninguém tinha pensado em usar esse sistema em outros organismos além das bactérias.

Outra descoberta básica, mas com diversos efeitos práticos obtida simultaneamente em três estudos (Sapranauskas et al., 2011; Gasiunas et al., 2012; Jinek et al., 2012), foi a determinação da atividade e especificidade dos dois domínios efetores na SpCas9. Foi determinada a função dos domínios HNH e RuvC presentes na SpCas9, sendo o primeiro responsável pela clivagem da fita complementar e o segundo responsável pela clivagem da fita não complementar de DNA (figura 1c).

3. Nickases

O conhecimento da especificidade desses domínios catalíticos da SpCas9 foi rapidamente explorado, e sistemas de edição genômica mais específicos foram criados utilizando versões modificadas da SpCas9 (Mali et al., 2013; Ran et al., 2013). Nesses estudos, os autores realizaram o detalhamento de variantes mutantes da SpCas9 contendo apenas um dos domínios catalíticos ativo. Essa estratégia criou duas **nickases**, i.e., versões da Cas9 capazes de reconhecer o DNA-alvo e realizar a clivagem de apenas uma fita do DNA. A primeira delas é denominada SpCas9 HNH⁺ RuvC⁻ (ou simplesmente **Cas9 HNH⁺**), na qual apenas o domínio catalítico HNH é funcional. Um exemplo muito conhecido desta *nickase* é a Cas9 D10A. A segunda, denominada SpCas9 HNH⁻ RuvC⁺ (ou simplesmente **Cas9 RuvC⁺**), apresenta apenas o domínio RuvC funcional, tendo como exemplo a Cas9 H840A. Os autores utilizaram essa característica

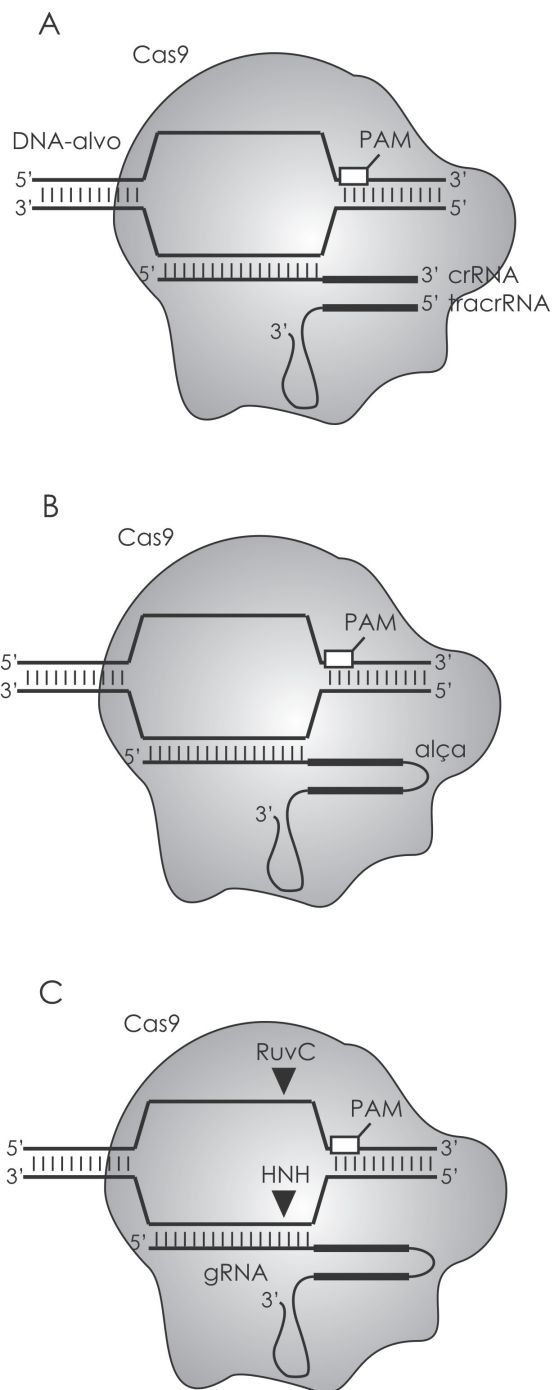


Figura 1. Modelo do sistema CRISPR-Cas9 em contato com DNA-alvo. A. Descrição do modelo natural contendo a Cas9, o DNA-alvo e os RNAs acessórios isolados (crRNA e tracrRNA). B. Adição da alça (*linker loop*) entre os dois RNAs acessórios. C. Descrição das regiões efetoras RuvC e HNH. Imagem: Pereira, TC (2016).

para criar um sistema mais eficaz e preciso de clivagem usando *pares da mesma nickase* (i.e., duas Cas9 HNH⁺ ou duas Cas9 RuvC⁺).

Nesse sistema é necessário o uso de dois sgRNAs, cada qual guiando uma *nickase* para cada uma das duas fitas de DNA. Nesse sentido, cada *nickase* realizaria um corte de fita simples. Consequentemente, a ação conjunta de ambas *nickases* resultaria na quebra da dupla fita (figura 2), que deve ser reparada pelo sistema de recombinação não homóloga (NHEJ), gerando assim o efeito desejado. O uso das *nickases* aumenta a especificidade de detecção do DNA-alvo (já que são envolvidos 2 sgRNAs) e diminui o efeito *off-target* em até 1.500 vezes, devido à necessidade da ocorrência de sequências específicas no DNA-alvo a uma distância de 23 nucleotídeos (sgRNA *offset*) (Ran et al., 2013).

4. Cas9 com ambos os sítios catalíticos inativos (*dead Cas9* - dCas9)

O conhecimento dos sítios catalíticos da SpCas9 e posterior capacidade de inativação destes propiciou a abertura de um novo leque de atividades para o sistema. A chamada **dCas9** (Cas9 *cataliticamente morta*) é capaz de reconhecer sequências específicas no genoma sem causar clivagem da sequência-alvo. Porém, os primeiros testes em bactérias mostraram que a mera associação do complexo dCas9 ao DNA-alvo levava à inibição da transcrição do gene-alvo, fosse pelo impedimento da associação da RNA polimerase ou de fatores de transcrição

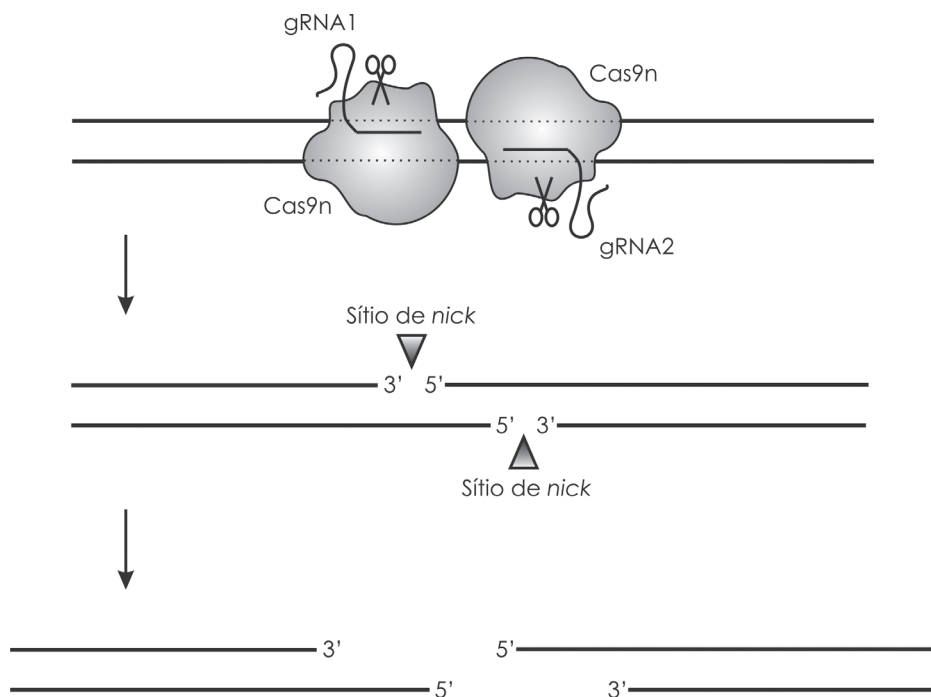


Figura 2. SpCas9 *nickase*. Modelo de ação dos pares de Cas9 *nickases*. A ação de cada *nickase* é dependente da existência de um gRNA específico. A quebra da dupla fita de DNA só ocorre na presença de duas *nickases* simultaneamente no mesmo loco. As tesouras representam domínios catalíticos (RuvC ou HNH). Imagem: Pereira, TC (2016).

(Qi et al., 2013). Nesse estudo, esse sistema foi chamado de **CRISPRi** (*CRISPR interference*, vide cap. 3), uma vez que o complexo dCas9/sgRNA poderia ser utilizado para bloqueio da expressão gênica em bactérias ou células eucarióticas (Bikard et al., 2013; Zhao et al., 2014).

Outras aplicações da dCas9 derivaram do fusão dela a diversas proteínas. Por exemplo, dCas9-GFP permite a marcação de DNA (*DNA labeling*, cap. 3), ao passo que dCas9-KRAB permite a regulação da expressão gênica (cap. 3).

5. Entendimento e desenvolvimento da região PAM

Desde a descoberta do sistema CRISPR com a Cas9 derivada de *S. pyogenes* foi verificada a exigência de uma sequência auxiliar de reconhecimento chamada PAM (figura 1a). Essa sequência, caracterizada por três nucleotídeos adjacentes ao DNA-alvo, é necessária para ação da SpCas9 e para o primeiro contato da nuclease com o DNA-alvo (vide cap. 2). No caso da SpCas9, as sequências de DNA-alvo são sempre flanqueadas pelo PAM composto pelos nucleotídeos “NGG” (Jinek et al., 2012), limitando o potencial uso da mesma apenas para sequências contendo esse PAM.

Em busca de alterar essa limitação do sistema, estudos baseados na estrutura molecular de SpCas9 buscaram modificar o sítio de reconhecimento PAM. Para isso foram utilizadas estratégias de *evolução direcionada* e *design combinatório de sequências*, a fim de mudar a especificidade da SpCas9 e permitir o reconhecimento de novas sequências PAM (Kleinstiver et al., 2015). Esta abordagem identificou versões mutantes da SpCas9 capazes de reconhecer outros PAMs como “NGA” ou “NGCG” em células de *zebrafish* ou humanas. Uma vez comprovada a possibilidade de variação do PAM baseado na estrutura da SpCas9, os autores verificaram que em outras espécies de bactérias, variantes ortólogas da SpCas9 como a **St1Cas9** (Deveau et al., 2008) encontrada em *Streptococcus thermophilus* e **SaCas9** (Ran et al., 2015) encontrada em *Staphylococcus aureus*, reconhecem naturalmente outros PAMs. Em ambos os casos, essas bactérias codificam proteínas menores comparadas à SpCas9 e seus PAMs foram determinados como “NNGGGT” para SaCas9 e “NNAGAA” para St1Cas9 (Kleinstiver et al., 2015). Outro exemplo é a **NmCas9**, derivada de *Neisseria meningitidis*, que reconhece um PAM mais extenso composto por 8 nucleotídeos “NNNNGATT” (Hou et al., 2013).

Esses trabalhos abriram uma nova fronteira para o uso do sistema CRISPR-Cas9 como modificador de ácidos nucleicos, uma vez que possibilita o uso de Cas9 específicas para o PAM encontrado na sequência-alvo desejada. Nesse sentido, diversos métodos foram desenvolvidos para a identificação/seleção de variantes de Cas9 com reconhecimento de PAMs alternativas.

O primeiro método é baseado na análise *in silico* dos fragmentos espaçadores do sistema CRISPR (Horvath et al., 2008; Mojica et al., 2009) e busca as sequências PAM mais conservadas entre os sítios-alvos, porém esse método é limitado pelas sequências depositadas em bancos de dados. Adicionalmente, ele pode conter PAMs mutados nas bactérias analisadas.

Métodos *in vivo* com o mesmo fim foram desenvolvidos. Nesse caso, bibliotecas de plasmídeos são desenhadas especificamente para o teste de todas as possíveis combinações de nucleotídeos encontrados na região PAM. Essa biblioteca é então clonada em uma região adjacente a um gene de resistência a antibiótico. Após o cultivo das bactérias contendo os plasmídeos da biblioteca, em meio de cultura com o antibiótico de seleção apenas aquelas que contêm um PAM inativo vão sobreviver. Já aquelas que contêm um PAM correto serão

destruídas, uma vez que o gene de resistência será inativado. Ao fim do experimento, todas as bactérias resistentes são sequenciadas, possibilitando a determinação das sequências PAM inativas. Desse modo, todas as outras sequências não obtidas no sequenciamento serão as necessárias para a ação da Cas9 em teste (Jiang et al., 2013; Karvelis et al., 2015).

Em uma tentativa de simplificar esta metodologia, foi desenvolvido um novo tipo de teste baseado em *seleção positiva*. Nesta abordagem, a confirmação da clivagem da PAM alternativa resulta na ativação de fluorescência (Leenay et al., 2016), facilitando o processo de seleção dos eventos positivos. A velocidade desse protocolo possibilitou a identificação de novos PAMs para diversas bactérias.

6. Cas9 de alta precisão

Até o momento relacionamos variantes de Cas9 utilizadas para diversos fins, sempre com o uso de sua especificidade em relação ao DNA-alvo. Porém, desde o princípio dos estudos com a SpCas9 foi verificado que esse sistema de reconhecimento é passível de erros (Cradick et al., 2013; Fu et al., 2013; Pattanayak et al., 2013), sendo que outras sequências de DNA não desejadas podem reconhecidas e clivadas. Essa característica de reconhecimento imperfeito é importante para o ambiente natural de ação da Cas9, onde ela tem como função controlar infecções por bacteriófagos. Nesse ambiente, a identificação de sequências ligeiramente diferentes facilita o combate de uma maior gama de invasores. Porém, com o desenvolvimento do uso desse sistema com a função de editar ácidos nucleicos, tanto para inativação gênica quanto para inserção de transgenes, foi notada a importância do aprimoramento do sistema.

Nesse sentido de reduzir o efeito *off-target*, o uso de *nickases* foi proposto, uma vez que elas precisam atuar aos pares para a produção de um evento de clivagem dupla fita (Mali et al., 2013; Ran et al., 2013). Como esta abordagem demanda dois sgRNAs para que o evento ocorra, o processo torna-se mais específico, pois dificilmente uma sequência não alvo (*off-target*) apresentará complementaridade aos dois sgRNAs simultaneamente (e uma distância de 23 nt entre eles). É também importante lembrar que eventuais *off-targets* com complementaridade a *apenas um* dos sgRNAs sofrerão corte em apenas uma das fitas (um *nick*), que será reparado pela célula sem ocorrer edição.

Porém, o grande avanço nesse campo veio com o desenvolvimento de versões fusionadas da dCas9. Nesse caso, a dCas9 foi combinada com a nuclease FokI (**fCas9**) gerando um complexo dependente de dois sgRNAs presentes nas fitas complementares do DNA-alvo, a uma distância específica de 15 ou 25 pb, possibilitando a dimerização das nucleases e clivagem do DNA-alvo (figura 3) (Guilinger et al., 2014; Tsai et al., 2014; Wyvekens et al., 2015). Nessa estratégia, os efeitos *off-target* são diminuídos, uma vez que são necessários dois sítios-alvos existentes a uma exata distância para que ocorra a dimerização das nucleases. Esse sistema teve um efeito *off-target* diminuído em 140 vezes comparado com a Cas9 selvagem e 4 vezes melhor do que as *nickases* discutidas acima.

Ainda em busca da *Cas9 perfeita*, dois outros grupos se basearam na estrutura molecular da SpCas9 procurando maneiras de reduzir o efeito *off-target*. O primeiro teve como base as propriedades energéticas da interação entre a SpCas9 e o DNA-alvo. Nesse contexto, foi verificado que existe uma “sobra” energética nessa interação que pode levar ao reconhecimento de DNAs-alvos com sequências não exatamente complementares ao sgRNA.

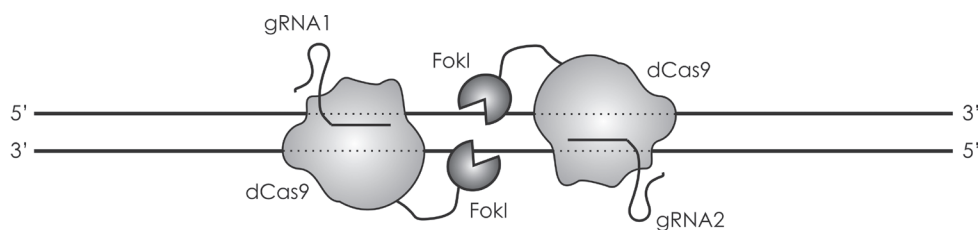


Figura 3. Esquema da estratégia dCas9 fusionado à nuclease FokI. Imagem: Pereira, TC (2016).

Com base nessa informação, foi realizado o desenho racional de diversas SpCas9 mutantes para quatro aminoácidos presentes na região de contato com o DNA-alvo e verificou-se que entre os 15 mutantes testados, 2 tinham mantido a capacidade de clivar alvos específicos e reduzido drasticamente os efeitos *off-target* (Kleinstiver et al., 2016). O produto desse estudo foi chamado **SpCas9-HF1** (*Streptococcus pyogenes* Cas9 High Fidelity 1), e diversos testes demonstraram que, quando usada *in vivo*, essa Cas9 mutante tem efeito *off-target* próximo a zero.

Simultaneamente, outro grupo descreveu a chamada **eSpCas9** (*enhanced specificity SpCas9*) (Slaymaker et al., 2016) também se utilizando da estrutura molecular do complexo DNA/SpCas9 (Anders et al., 2014; Nishimasu et al., 2014). Assim como no estudo anterior, foi proposto que existem posições na região de interação entre o DNA-alvo e o sítio de reconhecimento da Cas9 que favorecem a interação com alvos não específicos. A partir dessa hipótese, foram identificadas e testadas 31 posições passíveis de mutação, sendo que cinco combinações geraram uma diminuição dos efeitos *off-target*. Testes em escala genômica foram realizados e duas variáveis chamadas eSpCas9 1.0 e eSpCas9 1.1 foram selecionadas por terem função mantida e um nível de *off-target* próximo a zero (Slaymaker et al., 2016). Os dois trabalhos demonstram a plasticidade do sistema e sua capacidade de manter os efeitos desejados, mesmo após diversas mudanças nunca encontradas de forma natural.

7. Sistemas de ativação indireta da Cas9

Como discutido anteriormente, a definição da estrutura molecular do complexo DNA/SpCas9 foi muito importante para o desenvolvimento de novas versões da SpCas9. Outro aspecto derivado desse conhecimento foi o entendimento e demonstração de que a Cas9 pode ser dividida em duas subunidades não funcionais que, uma vez reunidas, podem executar sua função de clivagem do DNA-alvo (Wright et al., 2015). Essas subunidades foram chamadas de **Split-Cas9**. Com o uso dessa variante foram desenvolvidas diversas estratégias de ativação indireta do sistema, uma vez que as duas subunidades poderiam ser introduzidas nas células e posteriormente um sinal ativaria a aproximação destas.

Tendo como modelo as versões fusionadas da dCas9 discutidas anteriormente, os resultados mais promissores foram obtidos com Split-Cas9 fusionadas a proteínas “fotoativáveis” (gerando a *Photoactivatable Cas9*, **paCas9**). Em um exemplo desse método, cada subunidade foi fusionada a proteínas chamadas *pMagnet* e *mMagnet*, que sofrem dimerização quando excitadas por luz de comprimento de onda azul (412-426 nm) (Nihongaki

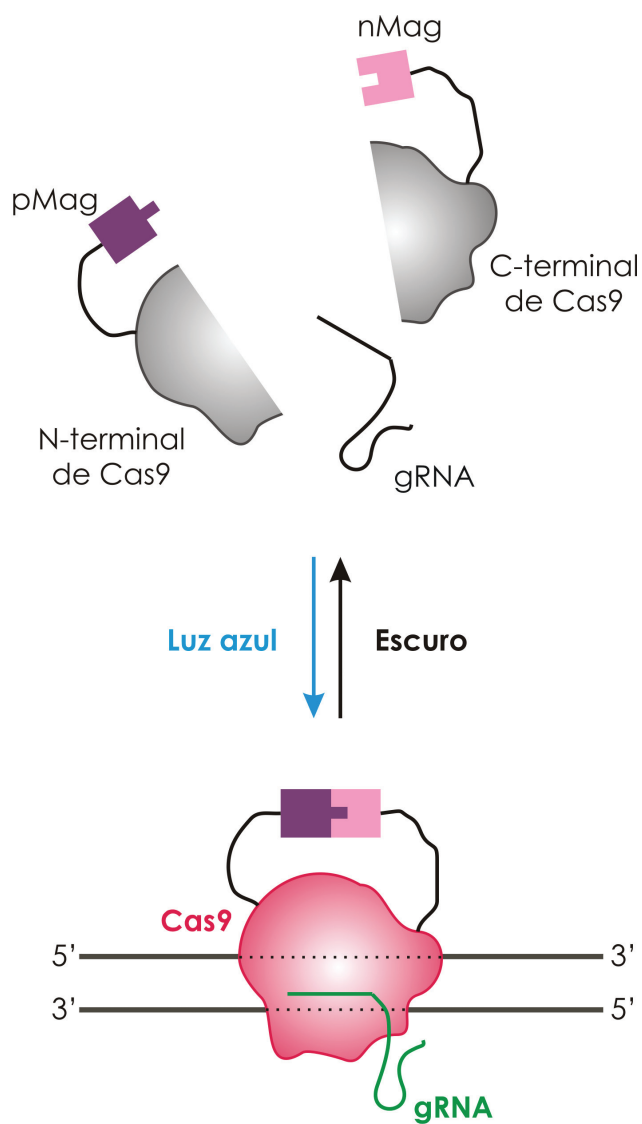


Figura 4. Esquema demonstrando as subunidades de Cas9 fusionadas ao sistema fotoativável. Esta estratégia pode ser aplicada tanto a Cas9 selvagem (para edição gênica) quanto dCas9 (para controle da expressão). Imagem: Pereira, TC (2016).

et al., 2015) (figura 4). Nesse contexto, células transfectadas com essas subunidades podem ser cultivadas, porém a ação da Cas9 só ocorrerá no momento desejado pelo pesquisador, possibilitando total controle do sistema.

Versões alternativas de controle de Split-dCas9 utilizaram outras proteínas com capacidade de dimerização controladas por drogas como a rapamicina, que induz a dimerização das proteínas FRB e FKBP (Zetsche, Volz, et al., 2015). Nesse caso, a dimerização e funcionamento do complexo só seria atingido com a administração da droga na cultura celular.

Em ambos os modelos, podemos notar que o sistema vem sendo desenvolvido e o controle de sua atividade está cada vez mais específico, possibilitando ao pesquisador o desenvolvimento de experimentos totalmente controlados.

8. Modificações nos sgRNAs

Até o momento discutimos diversas maneiras de melhorar os efeitos da técnica de CRISPR baseadas em alterações na nuclease Cas9. Uma outra estratégia para o mesmo fim é a alteração do sgRNA, uma vez que esse é necessário para o reconhecimento do DNA-alvo.

Os primeiros estudos nesse sentido foram realizados buscando a melhor combinação entre o tamanho do RNA-guia e o número de eventos *off-target*. De maneira contraintuitiva, foi observado que SpCas9 guiadas por sgRNAs truncados, contendo 17 nt, têm uma melhor performance gerando menos eventos *off-target* comparado ao sgRNA original de 20 nt (Fu et al., 2014). Uma possível explicação para isso seria a natureza promíscua dos últimos nucleotídeos da extremidade 5' da sequência-guia do gRNA, demonstrada por meio de ChIP-Seq com dCas9 (Kuscu et al., 2014). Portanto, a retirada destes últimos nt poderia explicar uma redução no número de *off-targets*.

Outro aspecto analisado do sgRNA foi sua longevidade. Pesquisadores notaram uma rápida degradação destas moléculas após a administração em células humanas primárias, como as células-tronco. Baseado em estudos com Interferência por RNA, já era conhecido que o uso de nucleotídeos modificados na produção de siRNAs aumentava a vida média destes. Com isso foram testadas diversas combinações de nucleotídeos na produção de sgRNAs. Os resultados mais expressivos foram obtidos com uso de uma combinação de 2'-fluoro (2'-F), 2'-O-metil (2'-O-Me) e *S-constrained ethyl* em posições específicas do sgRNA. Na combinação mais favorável foi verificada uma melhora expressiva na atividade do sistema comparado ao mesmo sgRNA produzido com nucleotídeos comuns (Rahdar et al., 2015). Outro estudo semelhante mostrou que o uso de sgRNAs sintéticos com nucleotídeos modificados produz melhores resultados quando comparados ao método usual, baseado em plasmídeos expressando os sgRNAs, especialmente em células humanas primárias como linfócitos T (Hendel et al., 2015).

9. Cpf1

A busca por novas possibilidades na área de edição genética continua. Recentemente um novo sistema CRISPR foi identificado na bactéria *Francisella novicida* a partir de análises de bioinformática, sendo classificado como um sistema CRISPR do Tipo V classe 2 (Makarova et al., 2015). A caracterização deste foi realizada pelo grupo do pesquisador Feng Zhang do

MIT, que evidenciou aspectos peculiares desse sistema não compartilhados pelos sistemas Cas9 codificados por outras bactérias.

O primeiro é a não existência de um tracrRNA no sistema. O segundo é a definição de uma sequência PAM rica em nucleotídeos “T”, especificamente “YTN”, diferente da encontrada no sistema Cas9 rica em “G”. A terceira é o método de clivagem do DNA: ele é mediado pela nuclease **Cpf1**, cujo corte da dupla fita de DNA gera pontas coesivas, diferente do sistema Cas9 que gera extremidades abruptas (*blunt end*). Outra distinção é a presença de dois sítios RuvC e a não existência do sítio HNH. Apesar das diferenças, o sistema CRISPR-Cpf1 tem capacidade de ação semelhante aos sistemas CRISPR-Cas9, interagindo com alvos diferentes devido a sua sequência PAM. Quando testado em células humanas, a Cpf1 é capaz de gerar efeitos semelhantes à Cas9, expandindo assim as possibilidades de edição do genoma (Zetsche, Gootenberg, et al., 2015).

10. Conclusões

Podemos apreciar no decorrer deste capítulo o rápido desenvolvimento e aprimoramento da tecnologia e da biologia relacionada ao sistema CRISPR-Cas9. Pesquisadores provenientes das mais variadas partes do mundo mostraram que é possível melhorar um sistema desenvolvido pela natureza desde o aparecimento das bactérias. Cada parte desse sistema foi analisado, decodificado e melhorado. Desde o desenho racional de estruturas da Cas9 até modificações na composição dos nucleotídeos dos sgRNAs. Essa determinação possibilitou o desenvolvimento de sistemas mais exatos, controlados e eficientes capazes de editar genomas de maneira específica e segura. Esses fatores culminaram na perspectiva do uso do sistema Cas9 em testes clínicos de fase 1 previstos para 2017, além da criação de diversas empresas focadas no uso clínico desse sistema. Desse modo, entendemos que o futuro da técnica de CRISPR deve nos surpreender, dada a velocidade do seu desenvolvimento.

Referências

- Anders C, et al. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. 2014 Sep;513(7519):569-73.
- Bikard D, et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*. 2013 Aug;41(15):7429-37.
- Cradick TJ, et al. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. 2013 Nov;41(20):9584-92.
- Deveau H, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol*. 2008 Feb;190(4):1390-400.
- Fu Y, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013 Sep;31(9):822-6.
- Fu Y, et al. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014 Mar;32(3):279-84.
- Gasiunas G, et al. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA*. 2012 Sep;109(39):E2579-86.
- Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 2014 Jun;32(6):577-82.
- Hendel A, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol*. 2015 Sep;33(9):985-9.
- Horvath P, et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol*. 2008 Feb;190(4):1401-12.
- Hou Z, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA*. 2013 Sep;110(39):15644-9.
- Jiang W, et al. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013 Mar;31(3):233-9.
- Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-21.
- Karvelis T, et al. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol*. 2015;16:253.

- Kleinstiver BP, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015 Jul;523(7561):481-5.
- Kleinstiver BP, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016 Jan;529(7587):490-5.
- Kuscu C, et al. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014 Jul;32(7):677-83.
- Leenay RT, et al. Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. *Mol Cell*. 2016 Apr;62(1):137-47.
- Makarova KS, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*. 2015 Nov;13(11):722-36.
- Mali P, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013 Sep;31(9):833-8.
- Mojica FJ, et al. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009;155(3):733-40.
- Nihongaki Y, et al. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat Biotechnol*. 2015 Jul;33(7):755-60.
- Nishimasu H, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014 Feb;156(5):935-49.
- Pattanayak V, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013 Sep;31(9):839-43.
- Qi LS, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013 Feb;152(5):1173-83.
- Rahdar M, et al. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci USA*. 2015 Dec;112(51):E7110-7.
- Ran FA, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013 Sep;154(6):1380-9.
- Ran FA, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015 Apr;520(7546):186-91.
- Sapranauskas R, et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res*. 2011 Nov;39(21):9275-82.
- Slaymaker IM, et al. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016 Jan;351(6268):84-8.
- Tsai SQ, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014 Jun;32(6):569-76.
- Wright AV, et al. Rational design of a split-Cas9 enzyme complex. *Proc Natl Acad Sci USA*. 2015 Mar;112(10):2984-9.
- Wyvekens N, et al. Dimeric CRISPR RNA-Guided FokI-dCas9 Nucleases Directed by Truncated gRNAs for Highly Specific Genome Editing. *Hum Gene Ther*. 2015 Jul;26(7):425-31.
- Zetsche B, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 2015 Oct;163(3):759-71.
- Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol*. 2015 Feb;33(2):139-42.
- Zhao Y, et al. Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. *Sci Rep*. 2014;4:3943.

Obtenção de Cas9 e gRNAs

Capítulo

7

Rodrigo Salazar da Silva^{1,a}, Erika Kague^{1,a}, Tiago Campos Pereira^{2,3,a},
Natalia Nardelli Gonçalves^{4,b} e Carlos Eduardo Ambrósio^{4,b}

¹*Dept.º de Genética e Biologia Evolutiva, Instituto de Biociências, USP, SP - Brasil*

²*Dept.º de Biologia, FFCLRP - USP, SP - Brasil*

³*Programa de Pós-Graduação em Genética, FMRP - USP, SP - Brasil*

⁴*Dept.º de Medicina Veterinária, FZEA, USP, SP - Brasil*

^a*Autoria das seções: 1-5 e 7*

^b*Autoria da seção: 6*

Estruturação do capítulo

1. Introdução
2. A nuclease Cas9
 - 2.1 Formatos disponíveis: vetores, mRNAs e proteínas
 - 2.2 Vias de obtenção: acadêmicas, comerciais e síntese
3. O gRNA
 - 3.1 Formatos disponíveis: vetores e RNAs
 - 3.2 Vias de obtenção: acadêmicas, comerciais e síntese
4. Busca por gRNAs indisponíveis comercialmente
5. Desenho de gRNAs
 - 5.1 Posição do gRNA em relação ao gene
 - 5.2 Outros aspectos
6. Construção de bibliotecas de gRNAs
 - 6.1 *Pooled libraries*
 - 6.2 Tipos de *pooled libraries*
7. Conclusões

1. Introdução

Após compreender os mecanismos moleculares básicos do funcionamento de CRISPR, sua versatilidade, aplicações e as variantes de Cas9 e gRNAs, o próximo passo é obter as moléculas para a realização dos experimentos. Este capítulo tem por finalidade apresentar as diversas formas comerciais, acadêmicas e de síntese em laboratório para obter os **reagentes de CRISPR** - uma das formas como são conhecidos Cas9 e o gRNA.

2. A nuclease Cas9

2.1 Formatos disponíveis: vetores, mRNAs e proteínas

A enzima Cas9 ou suas variantes (Cas9n, dCas9, etc.) podem ser administradas no modelo de estudo (célula/organismo) basicamente em uma de três formas: (i) DNA (vetor de expressão da Cas9), (ii) mRNA da Cas9 ou (iii) o próprio polipeptídeo da Cas9 (figura 1).

Os *vetores de expressão* (plasmidiais ou virais) são pequenas moléculas de dsDNA tipicamente circulares nas quais se encontra clonada a proteína de interesse, nesse caso a Cas9, sob o controle de um promotor e um terminador transcricionais específicos. Eles apresentam como vantagens: (i) a facilidade de se manipular em laboratório uma molécula estável (o DNA); (ii) possibilidade de se aumentar a massa disponível via amplificação em bactérias; (iii) estarem disponíveis em uma série de empresas, em diferentes configurações (diferentes promotores, variantes de Cas9, proteínas fusionadas e resistência a antibióticos para seleção); e (iv) maior estabilidade na célula. Geralmente os plasmídeos são introduzidos na célula/organismo via transfecção/eletroporação/infecção/injeção direta, para consequente transcrição/tradução *in vivo*. Os vetores podem ser singulares para a Cas9 ou ainda podem conter o sgRNA, para entrega simultânea dessas duas moléculas à célula. Contudo, uma vez que a expressão de Cas9 (assim como do gRNA) é dependente de promotores, que geralmente são espécie-específicos, observa-se um número limitado de espécies/grupos para os quais há vetores disponíveis (humano, rato, camundongo, monocotiledôneas, dicotiledôneas e alguns outros).

O *mRNA de Cas9*, por sua vez, está pronto para ser traduzido, podendo ser mais interessante em duas situações. A primeira se refere aos casos em que não há disponibilidade de vetores com promotores específicos para sua espécie; a introdução direta do mRNA elimina a dependência destas sequências regulatórias. A segunda se refere a situações em que o *tempo do experimento* é limitante. Por exemplo, trabalhos com zigotos e embriões (cuja janela experimental é curta) frequentemente se utilizam destas moléculas, pois elas precisam passar apenas pelo processo de tradução. Os vetores, em contraposição, precisam primeiro ser internalizados pelo núcleo, em seguida transcritos, posteriormente exportados para o citoplasma, para só então serem lidos pelos ribossomos. Entretanto, mRNAs são moléculas muito instáveis e, comparativamente, mais difíceis de serem manipuladas do que vetores de DNA.

Por fim, o *polipeptídeo de Cas9* não precisa passar por nenhum processo celular anterior, ele está apto para exercer sua atividade assim que entra no ambiente intracelular. Adicionalmente, a proteína Cas9 possui uma vantagem central: *ela é funcional em qualquer espécie*, sendo esta uma característica que não se aplica aos vetores e mRNAs de Cas9.

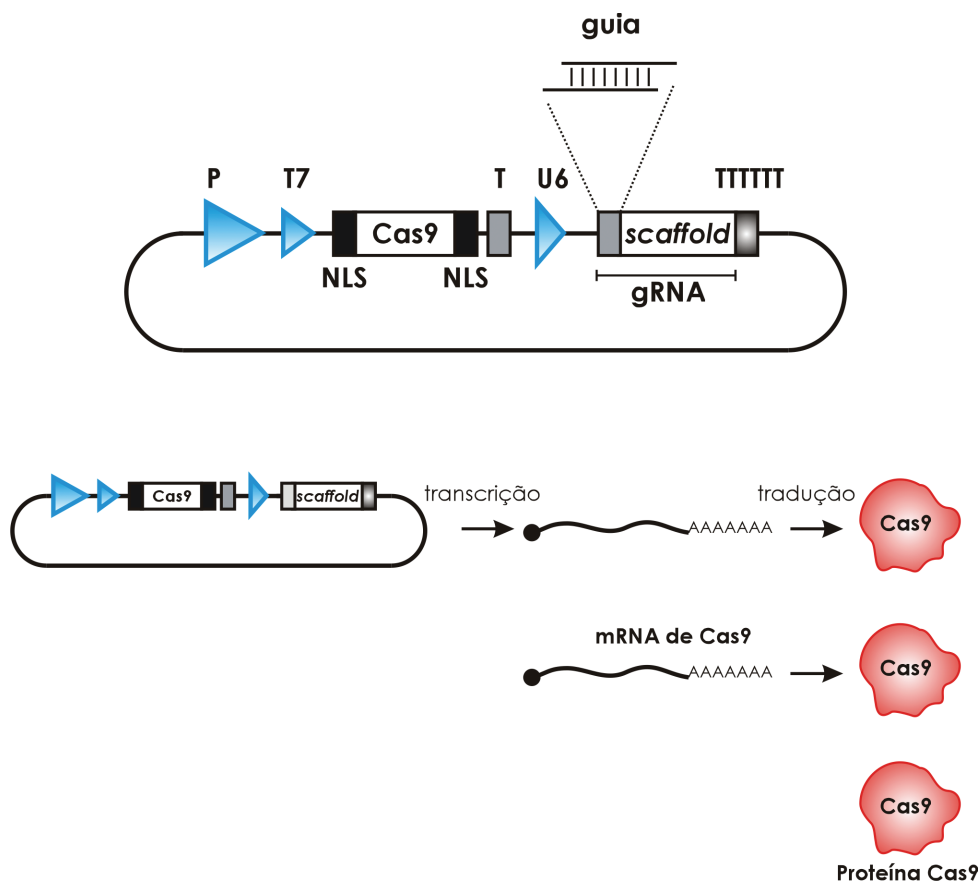


Figura 1. Formas de aquisição de Cas9 e gRNA. Plasmídeos expressando Cas9/gRNA são mais fáceis de se manipular em laboratório e estão disponíveis em diversas configurações, porém demandam diversas etapas de processamento intracelular. RNAs mensageiros de Cas9 são úteis em experimentos de breve janela temporal, mas a manipulação exige muitos cuidados. A proteína Cas9 é funcional em qualquer espécie, contudo apresenta elevado custo. P: promotor para Cas9. T7: promotor de origem viral, para transcrição *in vitro* de Cas9. NLS: sinal de localização nuclear. T: sinal de terminação da transcrição da Cas9. U6: promotor de RNA polimerase III para expressão do gRNA. Scaffold: sequência universal. "TTTTTT": sinal de terminação da transcrição de gRNA. Imagem: Pereira, TC (2016).

Para entender melhor isto, precisamos nos lembrar de dois aspectos: (i) o código genético é *quase* universal e (ii) espécies diferentes usam os códons para um determinado aminoácido com frequências distintas (fenômeno conhecido como *codon usage*). Portanto, uma sequência bacteriana (*e.g.*, o gene codificar da SpCas9) não necessariamente será traduzida de maneira correta e eficiente em células vegetais ou animais. Portanto, o uso de vetores ou mRNAs contendo a sequência nucleotídica de SpCas9 podem ser eficientes para a edição genética em diversas espécies de bactérias, mas menos eficientes em fungos, plantas e células humanas.

Foi pensando nisso que as empresas modificaram a sequência nucleotídica de SpCas9, tornando-a *otimizada para tradução* em células humanas (*human codon-optimized*, algumas vezes conhecida como **hCas9** ou **HsCas9**). Há versões otimizadas para milho, arroz, arabidopsis etc., muitas delas modificadas por grupos específicos de pesquisadores. Por um lado, essas otimizações aumentam as chances de sucesso da edição genética na espécie de interesse, mas considerando a miríade de espécies existentes, gerar versões otimizadas para cada uma delas

é algo impraticável. Nesse contexto mais amplo, a aquisição do *polipeptídeo da Cas9* emerge como uma solução universal simples. Sua grande desvantagem, no momento, é o preço.

Independentemente do tipo de molécula (DNA, RNA ou proteína), um elemento muito importante que deve estar integrado à Cas9 é o *sinal de localização nuclear (NLS)*, uma pequena sequência de resíduos de aminoácidos (~10 aa) com papel de direcionar a proteína para o núcleo. Esta pequena sequência (ou *motivo*) pode ser posicionada na extremidade amino (Wu et al., 2014) ou carboxi (Gao et al., 2015). Contudo, alguns autores relatam que o uso em apenas uma das duas extremidades não é tão eficiente quanto nas duas simultaneamente (Hou et al., 2013), fazendo com que muitos grupos utilizem versões NLS-Cas9-NLS (Platt et al., 2014; Yang et al., 2016). É essencial ressaltar que o NLS é fundamental quando se utiliza CRISPR em aplicações no DNA nuclear (edição genética, *DNA labeling*, regulação da expressão e mapeamento gênico) mas não no uso contra RNA citoplasmático (clivagem ou rastreamento de RNAs).

Como exemplos, vemos que em plantas a expressão de Cas9 geralmente se dá pela entrega de vetores, por meio de transformação via *Agrobacterium tumefaciens* (Ali et al., 2015) ou via bombardeamento de partículas (Shan et al., 2013). Já em camundongos há frequente microinjeção de plasmídeo (Kalebic et al., 2016) ou de mRNA (Chu et al., 2016), com alguns relatos de uso da forma proteica da Cas9 (Aida et al., 2015). Estudos com *Xenopus laevis* comumente se utilizam do mRNA (Blitz et al., 2013) ou de proteína Cas9 (Bhattacharya et al., 2015); em *zebrafish* a nuclease já foi entregue sob todas as formas, apresentando alta eficiência quando injetada em forma de proteína devido à grande estabilidade (Ablain et al., 2015; Varshney et al., 2015; Kotani et al., 2015). Já em modelos de cultura de células, a Cas9 tem sido entregue em forma de DNA (plasmídeos) e, em alguns casos, a proteína também foi utilizada (Park et al., 2015). A transfecção/eletroporação com mRNA parece ainda ser tóxica em culturas de células, e formas mais eficientes de entrega de mRNA na célula ainda estão sendo desenvolvidas.

2.2 Vias de obtenção: acadêmicas, comerciais e síntese

Acadêmica

Talvez o melhor exemplo de obtenção de vetores por meio acadêmico seja a Addgene, que é um repositório de plasmídeos depositados por pesquisadores, onde se encontram a maioria das construções já publicadas em revistas científicas (<http://www.addgene.org/CRISPR/>). A Addgene disponibiliza tanto a molécula do vetor em si quanto as informações de sua sequência nucleotídica (e o mapa da construção). Os plasmídeos podem ser buscados por grupo de pesquisa que os depositou, por tipo de função (quebra em dupla fita, quebra em apenas uma das fitas, domínios de ativação, purificação, visualização, etc.) e também por organismo modelo. Exemplos de plasmídeos mais utilizados (fornecidos pela Addgene) estão dispostos na Tabela 1.

Comercial

Alternativamente, uma gama muito extensa e interessante de plasmídeos é disponibilizada comercialmente por empresas como a Sigma-Aldrich, OriGene, IDT, GenScript, BioCat, ThermoFisher Scientific, ToolGene, PNAbio, New England BioLabs (vide Tabelas 2 e 3). Existe uma série grande e crescente de empresas dedicadas à comercialização de produtos

Tabela 1. Exemplos de plasmídeos frequentemente utilizados.

Plasmídeo	Característica	Organismos	Referências
hSpCas9-2A-Puro	Expressão de Cas9 e sítio de clonagem para sgRNA. Entrega de ambos componentes através de um único plasmídeo. Uso de puomicina, para selecionar células transfectadas. Há plasmídeos alternativos com Cas9 <i>nickase</i> e também com GFP.	Mamíferos	Ran et al., 2013
lentiCRISPR v2	Expressão de Cas9 e sgRNA em um único plasmídeo. Resistência a puomicina. Cas9 com NLS e humanizada.	Mamíferos	Sanjana et al., 2014
hCas9	Expressão de Cas9 humanizada.	Mamíferos	Mali et al., 2013
LSL-Cas9-Rosa26TV	Expressão de Cas9 códon-otimizada para humanos, dependente de Cre; com NSL; coexpressão de EGFP para seleção de células transfectadas com sucesso; camundongos podem ser selecionados positivamente com neomicina ou negativamente com DTA. Pode ser usado em conjunto com vários sistemas de entrega de sgRNA (viral ou não viral).	Mamíferos	Platt et al., 2014
pLSC-5	Vetor lentiviral para expressão de Cas9 em fragmentos; possui sítio de clonagem para sgRNA; uso de puomicina para seleção positiva de células transfectadas com sucesso.	Mamíferos	Zetsche et al., 2015
pU6-(BbsI)_CBh-Cas9-T2A-BFP-P2A-Ad4E1B	Expressão de Cas9 códon-otimizada para mamíferos, com NSL; possui sítio de clonagem para sgRNA; coexpressão de BFP para seleção de células transfectadas com sucesso.	Mamíferos	Chu et al., 2015
MSP469	Expressão de Cas9 variante VQR (reconhece PAMs NGAN e NGCG) códon-otimizada para mamíferos, com NSL; possui sítio de clonagem para sgRNA.	Mamíferos	Kleistiver et al., 2015
HBT-pcoCas9	Expressão de Cas9 códon-otimizada para plantas, com NSL; possui sítio de clonagem para sgRNA.	Plantas	Li et al., 2013
pRGER32	Vetor vazio usado para expressar Cas9 com promotor para ubiquitina de arroz e sgRNA com promotor U3 de arroz; possui sítio de clonagem para sgRNA; usado para plantas transformadas via <i>agrobacterium</i> ; seleção de plantas por higrômica.	Plantas	Xie et al., 2015
pME-Cas9	Expressão de Cas9 códon-otimizada para <i>zebrafish</i> , com NSLs; possui sítio de clonagem para sgRNA.	Zebrafish	Ablain et al., 2015
pCS2-Cas9 e cCS2+hspCas9	Expressão de Cas9 (original e forma humanizada).	Mamíferos Zebrafish Xenopus	Gagnon et al., 2014

Tabela 2. Opções de empresas que comercializam proteína e mRNA de Cas9 purificada.

	Nome	Nº catálogo proteína	Nº catálogo mRNA	Endereço	Descrição
Abm	Cas9 protein	K008 K030		https://www.abmgood.com/A330021E22Rik-CRISPR-Cas9-Proteins-A330021E22Rik.html	Oferece diversos tipos de proteínas Cas9 desde selvagem, <i>nickase</i> , NLS, desativada (<i>dead</i> Cas9).
New England BioLabs	Cas9 nuclease de <i>S. pyogenes</i> e Cas9 nuclease de <i>S. pyogenes</i> com NLS.	M0386S M0386T M0386L M0386M M0641S M0641T M0641L M0641M		https://www.neb.com/products/m0386-Cas9-nuclease-s-pyogenes#pd-properties-usage e https://www.neb.com/products/m0641-Cas9-nuclease-nls-s-pyogenes mRNA: não oferece	Oferece apenas um tipo de proteína Cas9. Produto em glicerol. A segunda enzima contém NLS (recomendada para uso, maior eficiência).
PNA Bio	Cas9	CP01 CP02 CN01 CD01	CR01 CR11	http://pnabio.com/products/Cas9.htm	Proteína Cas9 selvagem, <i>nickase</i> e desativada (<i>dead</i> Cas9) estão disponíveis. mRNA com NLS em forma selvagem e <i>nickase</i> .
Sigma	Cas9	TGEN-CP-50UG	213-791-2	http://www.sigmaaldrich.com/catalog/product/sigma/Cas9mrna?lang=pt&region=BR	Não apresenta descrição se NLS está presente.
System Biosciences	Transfection-ready Cas9 mRNA		CAS500A CAS502A CAS504A CAS530G CAS534G	https://www.systembio.com/CRISPR-Cas9-mrna-grna/ordering	Apresenta mRNA para Cas9 selvagem e <i>nickase</i> e outras versões com proteínas fluorescentes e também diferenciam em sistema procariótico e eucariótico.

Tabela 2. Continuação.

	Nome	Nº catálogo proteína	Nº catálogo mRNA	Endereço	Descrição
ThermoFisher Scientific	GeneArt™ Platinum™ Cas9 Nuclease;	B25641	A29378	https://www.thermofisher.com/order/catalog/product/B25641 e https://www.thermofisher.com/order/catalog/product/A29378?ICID=search-product	Apresenta NLS em proteína e mRNA.
	GeneArt™ CRISPR Nuclease mRNA	B25640			
	Recombinant Cas9 protein (<i>Streptococcus pyrogenes</i>); Cas9 mRNA	TGEN_CP1 TGEN_CP2 TGEN_CP3 TGEN_CP4	TGEN_CR1; TGEN_CN1	http://toolgen.com/?page_id=16 mRNA: http://toolgen.com/?page_id=16	
ToolGen					Empresa oferece diferentes tipos de proteínas de Cas9 (WT, <i>Nickase</i> e desativada) em diferentes concentrações. Produto liofilizado.
TriLink Biotechnologies	Cas9 mRNA	L-6116			Otimização para sistemas em mamíferos. RNAs com NLS, mímico de mRNA maduro e <i>nickase</i> .
		L-6125		http://www.trilinkbiotech.com/cart/scripts/prodList.asp?idCategory=196	
		L-6129			
		L-6325			

Tabela 3. Opções de empresas para obtenção comercial de sgRNAs.

Empresa	Endereço	O que oferecem
BioCat	http://www.biocat.com/functional-genomics	Biblioteca de sgRNAs disponível. Procura através de nome do gene. Espécies disponíveis varia entre genes. Fornecidos em plasmídeos. Clonagem customizada com fornecimento de sequência-alvo sem PAM.
GenScript	http://www.genscript.com/gRNA-database.html	Biblioteca de gRNA para humanos e camundongos, fornecidos em RNA ou clonado em plasmídeo. Oferecem mais de uma sequência-alvo para cada gene. Sistema de ativação gênica também disponível (SAM-Synergistic Activation Mediators).
GeneTools	http://toolgen.com/?page_id=254	Sequência ou nome do gene deve ser fornecida para pedir os sgRNAs. Fornecem sgRNAs clonados em plasmídeos, ou RNA e em diferentes concentrações. Possível pedir para diferentes organismos.
OriGen	http://www.origene.com/CRISPR-CAS9/Service.aspx	Necessário fornecer sequência-alvo de 100 pb ou as sequências-alvo já desenhadas. Pode ser escolhido o vetor para a clonagem do gRNA e concentração de produção.
PNABio	http://pnabio.com/support/gRNATool.htm order@pnabio.com .	Desenho a partir de região-alvo de menos de 1000 pb. Será gerada uma tabela com os possíveis alvos e inespecificidades. sgRNA pode ser pedido a partir de sequência-alvo ou região-alvo. Email deve ser enviado. Oferecem desenho de sgRNA para número relativamente grande de animais.
Sigma	http://www.sigmaaldrich.com/life-science/learning-center/CRISPR-request-form.html	Pedido customizado realizado através de formulário e fornecimento de número RefSeq do gene desejado. O organismo pode variar e deve ser verificado com a empresa. Bibliotecas são oferecidas para todo o genoma, com pelo menos dois alvos por gene (humano).
ThermoFisher	https://www.thermofisher.com/br/en/home/life-science/genome-editing/geneart-CRISPR/geneart-CRISPR-search-and-design-tool.html	Desenho de sgRNA. Biblioteca de sgRNA. Fornecido de diferentes formas e pareado com o tipo escolhido de Cas9 (plasmídeo, RNA, proteína). Escolha por nome de gene. Mais de um sgRNA por gene. Humano, camundongo e linhagens já prontas de células. Disponibilizam também <i>primers</i> para checagem da região-alvo para mutações. É necessário abrir conta e acessar o <i>cloud</i> deles.

relacionados a CRISPR. Atualmente, a ampla maioria deles oferece vetores, algumas disponibilizam o mRNA e poucas fornecem a proteína. Devido à crescente diversificação de produtos para CRISPR (variantes de Cas9, proteínas fusionadas, versões otimizadas para plantas e seres humanos, etc.) é importante ressaltar a necessidade de uma leitura cuidadosa de toda a descrição do produto com paciência antes de adquiri-lo. Listamos na Tabela 2 empresas que comercializam Cas9 em forma de mRNA e ou proteína.

Síntese

A síntese de vetores em laboratório é uma possibilidade adicional para grupos de pesquisa que possuem experiência em engenharia genética. Ela envolve a construção de um cassete de expressão contendo a Cas9 (com promotores e terminadores de sua espécie de interesse, fusionada ao NLS e com códons otimizados). A vantagem é que ele pode ser configurado a gosto.

A produção do mRNA de Cas9 em laboratório também é uma alternativa, envolvendo neste caso abordagens químicas (comercial/síntese). Esta estratégia envolve a aquisição comercial de um plasmídeo com Cas9 sob controle de um promotor viral T7 (T3 ou SP6). Esses plasmídeos podem ser submetidos a reação de transcrição *in vitro* em laboratório, utilizando kits comerciais com a RNA polimerase correspondente (*e.g.*, AmpliScribe™ T7 High Yield Transcription Kit, da Epicentre) e gerando imensas quantidades de mRNA de Cas9.

Por fim, a síntese do polipeptídeo da Cas9 em laboratório, via transcrição *in vitro* seguida por tradução *in vitro* (ou via expressão em bactérias e posterior purificação), apesar de possível não é usual, principalmente devido a questões de pureza.

3. O gRNA

3.1 Formatos disponíveis: vetores e RNAs

O gRNA pode ser entregue na célula/organismo de estudo em forma de DNA (vetores) ou a própria molécula de RNA. Ao contrário da Cas9, não há aqui questões relacionadas à otimização de códons nem NLS. Contudo, como mencionado anteriormente, a manipulação do DNA é mais fácil, enquanto o RNA é interessante em experimentos com janelas experimentais breves.

3.2 Vias de obtenção: acadêmicas, comerciais e síntese

Acadêmica

Novamente, o melhor exemplo de obtenção por meio acadêmico provavelmente é a Addgene, que também possui vetores com gRNAs clonados. Contudo, apesar da Addgene conter em seu banco uma ampla representatividade quanto às variantes de Cas9, o mesmo não ocorre, neste momento, quanto aos gRNAs/genes-alvo. Pode-se encontrar vetores com gRNAs para vários genes humanos e de camundongos, mas obviamente não para todos.

Comercial

Muitas empresas disponibilizam vetores com o gene da Cas9 e a *sequência universal* do gRNA (de ~80 nt, na extremidade 3', vide cap. 2). Próximo a esta sequência universal existe um sítio de clonagem, no qual é possível inserir a *sequência-guia* contra o gene-alvo de interesse (20 nt). Esta etapa resulta na *reconstrução do gRNA funcional* no vetor, podendo agora ser utilizado para os experimentos (figura 1). A vantagem destes vetores é que eles podem ser modificados a fim de se editar qualquer gene de interesse. A *sequência-guia* (~20 nt) pode ser adquirida como oligos de DNA (semelhantes a *primers*) que devem ser clonados no vetor para posterior uso no experimento.

Alternativamente, algumas empresas possuem *bancos especiais*, com vetores prontos para a edição gênica. São plasmídeos/vetores virais construídos já com a Cas9 e o gRNA funcional para seu gene de interesse, tornando-se desnecessário qualquer tipo de clonagem. Uma vez recebido no laboratório, o vetor pode ser prontamente utilizado no experimento. Apesar de extremamente práticos, há algumas limitações como o preço e disponibilidade para apenas algumas espécies (geralmente humano, rato e camundongo) e/ou genes. Por fim, ao contrário do exemplo do parágrafo anterior, estes vetores são exclusivos para um gene-alvo, não sendo facilmente modificados (*i.e.*, troca da sequência-alvo) para editar outros alvos.

Mais especificamente, a empresa BioCat (<http://www.biocat.com>) fornece os sgRNAs em forma de plasmídeos para humano, camundongo e rato, mas pode ser possível que outras espécies apareçam no resultado. Já pela GenScript é possível comprar sgRNAs em forma de RNAs, porém apenas está disponível para humano e camundongo. Dessa forma, estas companhias oferecem sgRNAs pré-desenhados que podem ser buscados pelo site a partir do nome do gene. Outras empresas fornecem o mesmo serviço, mas atualmente a lista de sgRNAs será enviada somente após contato por e-mail, entre elas a PNABio (<http://pnabio.com>) e ToolGene (<http://toolgen.com>). A limitação em regiões-alvo e organismos disponíveis que essas empresas fornecem ainda é grande. Uma forma de contornar este problema é trabalhar com empresas em que a sequência-alvo pode ser depositada, e eles trabalham em achar possíveis sítios-alvo e sintetizar seu sgRNA. Entre as empresas que trabalham dessa forma estão: Sigma (<http://CRISPR.sigmainformatics.com>), ThermoFisher Scientific (atualmente é preciso abrir uma conta) e ToolGene. Listamos as empresas que trabalham com sgRNAs na Tabela 3.

Síntese

A produção de vetores para a expressão de gRNAs em laboratório não é complexa. Ela envolve a construção de um cassete de expressão contendo, basicamente, três elementos. Primeiro, um promotor da RNA polimerase III (*e.g.*, U6) para expressão de pequenos RNAs, espécie-específico (mas não obrigatoriamente). Segundo, o gRNA em si, composto pela sequência-guia (~20 nt) seguida pela sequência universal (~80 nt). Por fim, o sinal de terminação da transcrição, que para Pol. III é tipicamente uma sequência de seis timinas ("TTTTTT") (figura 1).

O gRNA também pode ser obtido na forma de molécula RNA em si, via transcrição *in vitro* a partir do promotor T7 do plasmídeo, de forma semelhante ao descrito para mRNA de Cas9 (figura 1). As vantagens e desvantagens do uso deste formato são os mesmos discutidos anteriormente para o mRNA de Cas9.

4. Busca por gRNAs indisponíveis comercialmente

Novamente, a limitação em número de espécies e sequências-alvo presentes em bancos de sgRNA comercializados pelas empresas é um problema para aqueles que desejam trabalhar com outros organismos além de células humanas, camundongo e rato.

Então, como conseguir tais sgRNAs? Uma forma simples é pesquisar na literatura científica (*e.g.*, PUBMED - <http://www.ncbi.nlm.nih.gov/pubmed>) se há trabalhos com CRISPR direcionado contra seu gene de interesse em sua espécie de pesquisa (ou em alguma espécie próxima). Caso encontre algum artigo que tenha publicado CRISPR com o gene e organismo de interesse, contatar o laboratório pode ser válido para aquisição do vetor. Porém, como a tecnologia é relativamente recente, a chance de se obter um sgRNA dessa forma é baixa.

Como dito anteriormente, outra forma mais eficaz é obter a sequência nucleotídica do gRNA disponibilizada no artigo e seguir com um pedido customizado para as empresas. Algumas dessas companhias são mais versáteis e permitem a inserção da sequência nucleotídica do gene inteiro. Em seguida, elas mesmas desenham e sintetizam o gRNA (clonado em vetor ou já no formato de RNA).

5. Desenho de gRNAs

Como observado, há inúmeras formas de se obter o gRNA. Algumas muito práticas, como a aquisição comercial do plasmídeo com o gRNA clonado diretamente da empresa, pronto para o experimento. Em outros casos, ao informar à empresa o gene-alvo, ela mesma fica responsável por desenhar e sintetizar o gRNA.

Entretanto, em alguns casos, o pesquisador possui em seu laboratório um vetor com Cas9 e o *scaffold* do gRNA (a sequência universal de ~80 nt), necessitando apenas da *sequência-guia* (SG, de ~20 nt) para clonar no mesmo e iniciar seus experimentos. Caso esta SG não esteja disponível na literatura, pode-se desenhá-la de maneira bem simples, como veremos a seguir (revisto em Lee et al., 2016; revisto em Mohr et al., 2016).

5.1 Posição do gRNA em relação ao gene

Um aspecto-chave é a posição da SG em relação à estrutura do gene como um todo. Experimentos visando o nocauteamento de genes codificadores de proteínas devem dar preferência por sítios exônicos localizados logo após o códon de iniciação (*start codon*) (Figura 2). Uma vez que *indels* frequentemente alteram a fase de leitura, quanto mais próxima a mutação estiver da extremidade aminoterminal da proteína, maior será o número de aminoácidos modificados e, conseqüentemente, maior será a garantia de nocauteamento. Apesar da região promotora, 5' UTR, íntrons e 3' UTR poderem ser utilizados (observando-se certas condições), a preferência deve ser dada para a região codificadora.

Por sua vez, abordagens para *RNA tracking* e *clivagem de RNA* devem ser focadas apenas em éxons; para *DNA labeling*, em qualquer parte do gene. Já para ensaios de controle da expressão gênica (ativação ou repressão), os gRNAs devem ser direcionados às regiões promotoras (mas não exclusivamente) (Larson et al., 2013).

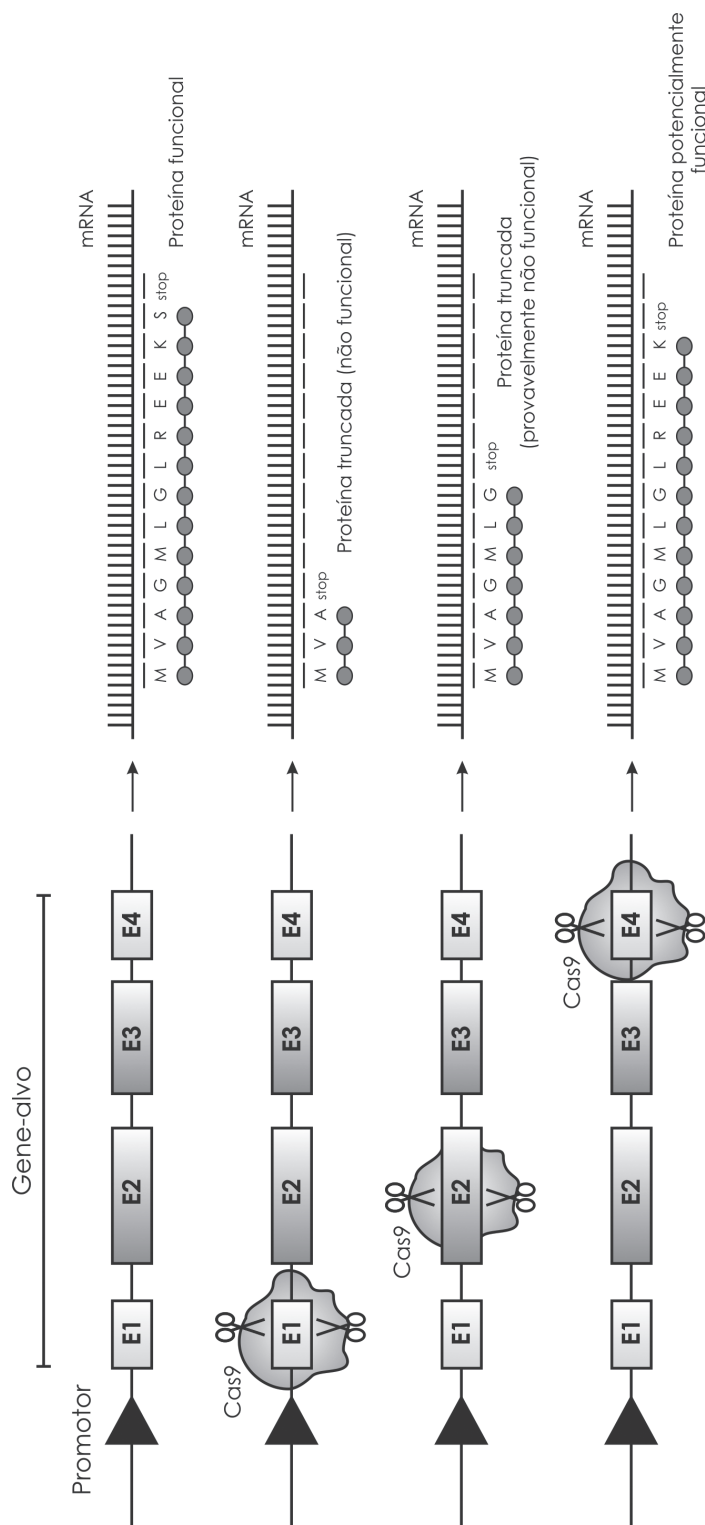


Figura 2. Relação entre o local de edição genética e o efeito sobre a proteína. E1-4: éxons de um a quatro. Letras: diferentes aminoácidos. Imagem: Pereira, TC (2016).

5.2 Outros aspectos

A forma mais direta de desenho de gRNAs individuais é por meio do uso de *softwares online*. Tais programas, listados na Tabela 4, consideram alguns parâmetros básicos de *design*, geram respostas rápidas e frequentemente satisfatórias para o sucesso do experimento.

Pesquisadores mais criteriosos podem desenhar *manualmente* a SG. Apesar de mais laboriosa, esta estratégia permite o total controle de todas as etapas do processo, algo que muitas vezes não é possível ao se utilizar *softwares*. As regras gerais para desenho são relativamente simples: uma sequência de vinte nucleotídeos (N_{20}) que deve ser *única no genoma* e apresentar um conteúdo “GC” próximo a 50% (Wang et al., 2014). Esta SG deve estar posicionada exatamente ao lado de uma PAM, que para a SpCas9 é o triplete “NGG”. Devido a isto, muitas vezes a expressão “ N_{20} NGG” é mencionada nos programas de busca por alvos.

Apesar de a SpCas9 ser a enzima mais comumente utilizada, existem outras nucleases cujas PAMs são diferentes. Por exemplo: Cas9 de *Neisseria meningitidis* (NmCas9) reconhece a PAM “NNNGATT” (Hou et al., 2013), de *Streptococcus thermophilus* (St1Cas9) “NNAGAAW” (Garneau et al., 2010), de *Staphylococcus aureus* (SaCas9) “NNGRR” (Ran et al., 2015). Portanto, este ajuste na procura por alvos deve ser considerado ao se utilizar outras Cas9 (e.g., “ N_{20} NNGRR” para SaCas9).

Alguns aspectos mais específicos adicionais devem ser considerados. Por exemplo, deve-se evitar que SG apresente sítios de restrição envolvidos no processo de clonagem do próprio cassete de expressão do gRNA. Segundo, frequentemente o gRNA está sob controle do promotor U6 (ou H1) da RNA polimerase III, enzima específica para transcrição de pequenos RNAs. Nesse sentido, a SG não pode conter quatro ou mais timinas (“TTTT”), pois este é o sinal de terminação desta polimerase (Mohr et al., 2016). Deve-se também evitar sítios que apresentem SNPs, pois isto pode impedir a perfeita interação entre SG e DNA-alvo. Por fim, deve-se evitar estruturas secundárias rígidas (e.g., *hairpins* ou *bulge loops* com temperaturas de *melting* acima de 5°C), aspecto que pode ser averiguado em programas como *Gene Runner* (www.generunner.net).

Uma curiosa etapa adicional é a avaliação do potencial de indução de mutações *frameshift* (Bae et al., 2014), útil para experimentos que objetivam o nocauteamento de genes codificadores de proteínas. Em alguns casos, o sítio de clivagem pode apresentar micro-homologia entre as extremidades de DNA geradas potencialmente restaurando a fase

Tabela 4. Relação dos sites mais utilizados para o desenho de sequências-guias que avaliam *off-targets* e apresentam interface web.

Website	URL
CRISPR design	http://crispr.mit.edu
E-CRISP	http://www.e-crisp.org/E-CRISP/
flyCRISPR	http://flycrispr.molbio.wisc.edu/tools
DRSC CRISPR	http://www.flyrnai.org/crispr2/
CRISPy	http://staff.biosustain.dtu.dk/laeb/crispy/
GT-Scan	http://gt-scan.braembl.org.au/gt-scan/
CHOPCHOP	https://chopchop.rc.fas.harvard.edu/
CLD	https://github.com/boutroslob/cld

de leitura e, conseqüentemente, impossibilitando o nocauteamento. Bae e colaboradores desenvolveram um *software* capaz de verificar o potencial da SG em gerar mutações de *frameshift*, que são as mais interessantes para nocaute.

A avaliação de especificidade da SG, *i.e.*, a busca por possíveis *off-targets*, pode ser feita por meio de diversos *softwares*, listados e detalhados por Lee e colaboradores (Lee et al., 2016). Uma alternativa é o popular programa BLASTn, (NCBI), no qual basta ajustar os parâmetros (*Algorithm parameters* – no fim da página): selecionar “*Short Queries*” e alterar “*Word Size*” para 20. Em princípio, SG não deveria apresentar quaisquer semelhanças com outra sequência no genoma, além do próprio DNA-alvo. Contudo, nem sempre isso é possível e sequências (*hits*) com algum grau de similaridade são comuns.

Neste momento devemos nos lembrar que a SG apresenta regiões aparentemente promíscuas e regiões de alta fidelidade (Hsu et al., 2013; Pattanayak et al., 2013; Fu et al., 2013). Estudos de ChIP-Seq com dCas9 revelaram que os nove nucleotídeos da extremidade 5' toleram mal pareamento com o DNA (Kuscu et al., 2014), ao passo que a região *seed* precisaria ser totalmente complementar (figura 3).

Portanto, caso a SG apresente similaridade com um alvo não desejado, mas ela difira na porção *seed*, isso sugere que este *off-target* não será editado acidentalmente, já que não haverá pareamento entre o gRNA e o DNA-alvo. Entretanto, caso a SG apresente similaridade

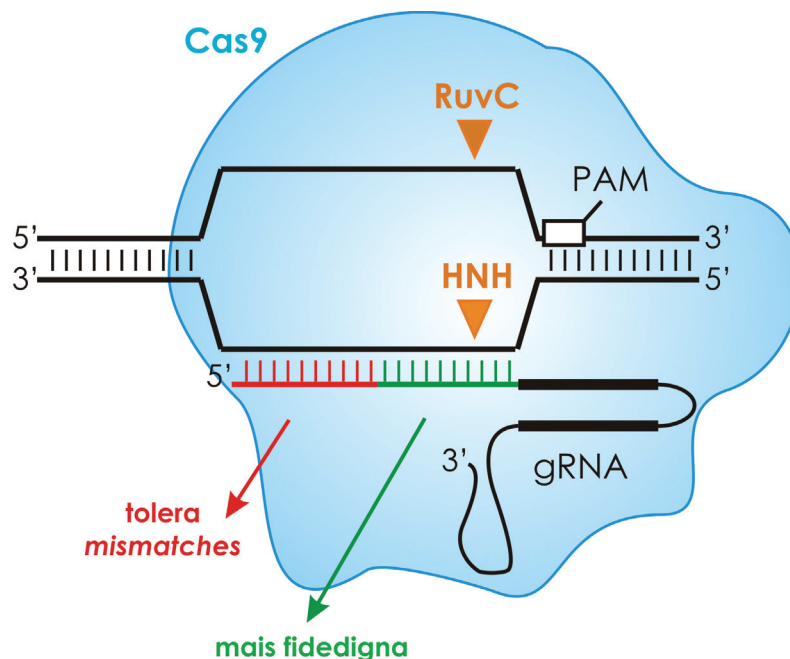


Figura 3. Especificidade da sequência-guia em função da posição dos nucleotídeos. Estudos têm evidenciado que a extremidade 5' do guia tolera mal pareamentos com o DNA-alvo, explicando a existência de *off-targets*. De forma oposta, a região mais próxima à 3' (*seed*) seria menos permissiva a pareamentos errôneos. Imagem: Pereira, TC (2016).

com um alvo não desejado, diferindo apenas na extremidade 5', isso sugere que este *off-target* poderá sim ser editado acidentalmente, já que a 5' naturalmente é promíscua.

Obviamente, a decisão quanto a situações intermediárias é muito difícil. Devo ou não utilizar uma SG cuja *seed* difira em apenas 1-2 nt de um *off-target*? Infelizmente nem todos os parâmetros que regem a interação entre SG e o alvo são conhecidos (Mohr et al., 2016). Mesmo os intrigantes dados de ChIP-Seq com dCas9 sobre a promiscuidade da 5' são questionáveis (Tsai et al., 2015; revisto em Lee et al., 2016). A sugestão é procurar por outras SGs e utilizar aquela que apresente menor similaridade. Adicionalmente, a análise molecular pós-experimento para verificar se houve ou não edição do potencial *off-target* identificado *in silico* é plenamente válida.

Existem trabalhos analisando o efeito da composição nucleotídica ao longo de toda a extensão da SG, *i.e.*, quais nucleotídeos devem ser *preferidos* ou *preteridos* em determinada posição (Doench et al., 2014), mas a opção pelo uso de parâmetros tão restritivos deve ser ponderada pelo leitor.

Outro aspecto muito importante é determinar se o gene-alvo apresenta formas alternativas de *splicing*, se possui outras cópias no genoma (parálogos) e se a intenção é editar apenas o alvo ou também estas outras sequências relacionadas. Tudo isso demandará análises mais profundas para encontrar uma sequência-alvo exclusiva para uma isoforma ou comum a todos eles.

Por fim, em situações atípicas, de genes sobrepostos por exemplo, a edição da sequência nucleotídica afeta ambas proteínas simultaneamente. Este desafio pode ser resolvido de uma forma elegante: clivagem do DNA-alvo e uso de um DNA doador cuja sequência tenha sido detalhadamente redesenhada, introduzindo mutações que sejam silenciosas ou neutras para uma proteína e, simultaneamente, *missense*, *nonsense* ou de *frameshift* para a segunda proteína.

6. Construção de bibliotecas de gRNAs

Um dos principais objetivos da genética, em uma visão mais ampla, é conhecer a função de cada um dos milhares de genes de cada um dos milhões de espécies no planeta. Técnicas de *genética reversa*, tais como RNAi e CRISPR, têm nos permitido avançar em direção a este objetivo maior. Contudo, adaptações para utilizar essas técnicas em larga escala tornam-se necessárias, caso queiramos alcançar este objetivo em um tempo razoável.

A construção de bibliotecas de gRNAs para todos os genes de uma determinada espécie constituem um desses aprimoramentos. Por meio delas somos capazes de nocautear e em seguida identificar o papel de cada um dos genes na célula/organismo. Exemplos recentes desse avanço compreendem bibliotecas para estudos em *Saccharomyces cerevisiae* (Ryan e Cate, 2014), drosófila (Bassett et al., 2015), células de camundongos (Koike-Yusa et al., 2014), genes humanos (Wang et al., 2014; Shalem et al., 2014; Zhou et al., 2014) e proteínas envolvidas na entrada e transmissão célula a célula do vírus da Hepatite C (HCV) (Ren et al., 2015). Adicionalmente, bibliotecas com mais de um gRNA para o mesmo gene-alvo permitem a busca por aqueles que são totalmente alvo-específicos, minimizando assim eventuais efeitos *off-targets*.

As bibliotecas podem ser divididas em *arrayed libraries* (bibliotecas dispostas em arranjos fixos) e *pooled libraries* (bibliotecas dispostas em misturas, em uma tradução livre).

Nas primeiras, os gRNAs são posicionados em placas de multi-poços contendo células, e cada poço recebe um gRNA conhecido a ser avaliado. Assim, a condução de tais experimentos exigem recursos especiais e robôs para manipulação de muitas placas. Porém, como vantagem, uma quantidade muito maior de fenótipos pode ser investigada usando, por exemplo, marcação por fluorescência ou abordagens baseadas em coleta automatizada de imagens. Protocolos simples e eficazes para construção de tais bibliotecas são descritos na revisão de Agrotis e Keteller (2015).

Já a construção de *pooled libraries* é comparativamente bem mais comum, simples e barata, sendo, portanto, o tema desta seção.

6.1 *Pooled libraries* – visão geral

Aspectos gerais

Os princípios básicos para construção de uma *pooled library* envolvem: (i) obtenção *in silico* das *sequências-guias* (i.e., os 20 nt do gRNA que direcionam a Cas9), (ii) síntese dos oligonucleotídeos correspondentes, (iii) ligação ao vetor, (iv) transfecção/transdução em condições para garantir um vetor por célula, (v) isolamento dos clones e (vi) checagem da biblioteca (sequenciamento dos clones) (Shalem et al., 2015).

Obtenção in silico das sequências-guias

De acordo com Wang et al. (2016), um primeiro e importante passo é a obtenção da lista de *sequências-guias* para o conjunto de genes ou genoma de interesse. No caso das espécies humana e murina, este grupo de pesquisa gerou uma lista de sgRNAs previamente validados que estão disponíveis gratuitamente em <http://www.broadinstitute.org/~timw/CRISPR/> para síntese.

Nesse sentido, quando se deseja construir bibliotecas pequenas, abrangendo poucos genes humanos ou murinos, pode-se usar a lista anteriormente mencionada ou uma das muitas ferramentas *online* disponíveis para desenho de *sequências-guias* (Tabela 4). Recomenda-se o desenho de pelo menos três *sequências-guias* para minimizar efeitos *off-targets* (Malina et al., 2014). O conteúdo GC deve ser equilibrado: quantidades muito altas ou muito baixas devem ser evitadas (Wang et al., 2014).

Contudo, quando se objetiva desenhar guias para todo um genoma, estratégias mais sofisticadas são necessárias. Um exemplo de tais ferramentas de bioinformática foi descrito por Heigwer e colaboradores (2016), com um método integrado para customizar bibliotecas de sgRNAs para qualquer organismo que tenha seu genoma sequenciado disponível. Além de ser compatível com o uso convencionais de CRISPR (clivagem de genes codificadores de proteínas), o recurso CLD (*CRISPR library designer*) também é compatível para o uso de Cas9 modificadas, assim como alvos não codificadores de proteínas. Para demonstrar a eficácia de tal ferramenta foi feito um escaneamento para a via TRAIL (*TNF-related apoptosis inducing ligand*) usando uma biblioteca customizada de 12.471 sgRNAs. O CLD funciona conforme os seguintes passos: (i) obtenção da sequência ENSEMBL, lista dos genes e os parâmetros que devem ser usados (tamanho mínimo do desenho, máximo, porcentagem de “GC”); (ii) índice das sequências genômicas, gerado por uma previsão e filtragem dos sítios de sgRNAs-alvos a partir da lista de genes; (iii) avaliação e remoção de possíveis *off-targets*; (iv) um arquivo com

os melhores sgRNAs para cada gene; (v) um arquivo final de saída, incluindo o formato FASTA com as sequências de oligonucleotídeos para síntese e posterior clonagem no vetor-alvo.

Clonagem por digestão e ligação

Após a seleção *in silico* das sequências-guias e a compra dos oligonucleotídeos correspondentes (da mesma forma que se adquire *primers*), estes devem ser clonados no vetor. Há várias formas para isso, abordaremos aqui duas delas. A primeira envolve a digestão do vetor e ligação direta dos oligonucleotídeos (duplexes já anelados), caso eles tenham sido desenhados de tal forma que cada duplex já apresente as extremidades coesivas compatíveis ao vetor digerido.

Alternativamente, os oligonucleotídeos podem ser submetidos a PCR, seja para *amplificar* o número de moléculas, seja para *anexar alguma sequência* a eles (como o promotor U6 ou o *scaffold*). De qualquer forma, após este procedimento os amplicons não apresentarão mais as extremidades coesivas, sendo necessário então a digestão deles com a mesma enzima de restrição usada no vetor.

Após purificação com kits comerciais, os amplicons podem ser clonados em vetor contendo Cas9 (figura 4), que deve ser escolhido de acordo com o foco desejado. A empresa Addgene, por exemplo, disponibiliza uma ampla série de plasmídeos com Cas9 (<https://www.addgene.org/crispr/zhang/>) para a clonagem dos gRNAs. Esta grande diversidade de vetores permite a construção de bibliotecas para mamíferos, bactérias, leveduras, *C. elegans*, drosófila, *zebrafish*, *Xenopus* e plantas. Lembrando que outras empresas também comercializam diversos outros plasmídeos igualmente interessantes.

O vetor é então internalizado em bactérias competentes por protocolos convencionais, como eletroporação/transformação, as colônias positivas selecionadas são purificadas, o DNA é extraído e sequenciado na região do sgRNA. A biblioteca resultante deste procedimento pode ser usada para a transfecção/transdução (no caso de vírus) das células por protocolos convencionais mais adequados (*e.g.*, vetores lentivirais para células 293).

Clonagem por LIC (ligation-independent cloning)

A clonagem do oligonucleotídeo ao vetor pode também ser feita via método LIC (*ligation-independent cloning*), que não é novo no ramo das tecnologias de edição gênica, já tendo sido empregado para estudos com TALENS (Schmid-Burgk et al., 2013). Esta mesma abordagem foi recentemente aplicada por Schmidt e colaboradores (2015), na qual uma biblioteca com 19.506 sgRNAs únicos e validados foi testada em células 293, com 100% de *knockout* para os 19 genes avaliados.

Resumidamente, o sistema funciona a partir de um plasmídeo linearizado com as enzimas *ApaI* e *SpeI*. Em seguida, com uso da T4 DNA polimerase e dTTP em um processo especial denominado *chewback reaction*, as extremidades do vetor sofrem recessão, permitindo um extenso pareamento do oligonucleotídeo com o vetor. Mesmo não estando covalentemente ligadas, estas moléculas (vetor/oligonucleotídeo) podem ser introduzidas em bactérias competentes (via transformação), sendo então ligadas pela maquinaria de reparo da bactéria. Note, portanto, que diferentemente de um processo comum de clonagem com restrição-ligação *in vitro*, LIC provê uma rota eficiente para ligação *in vivo*.

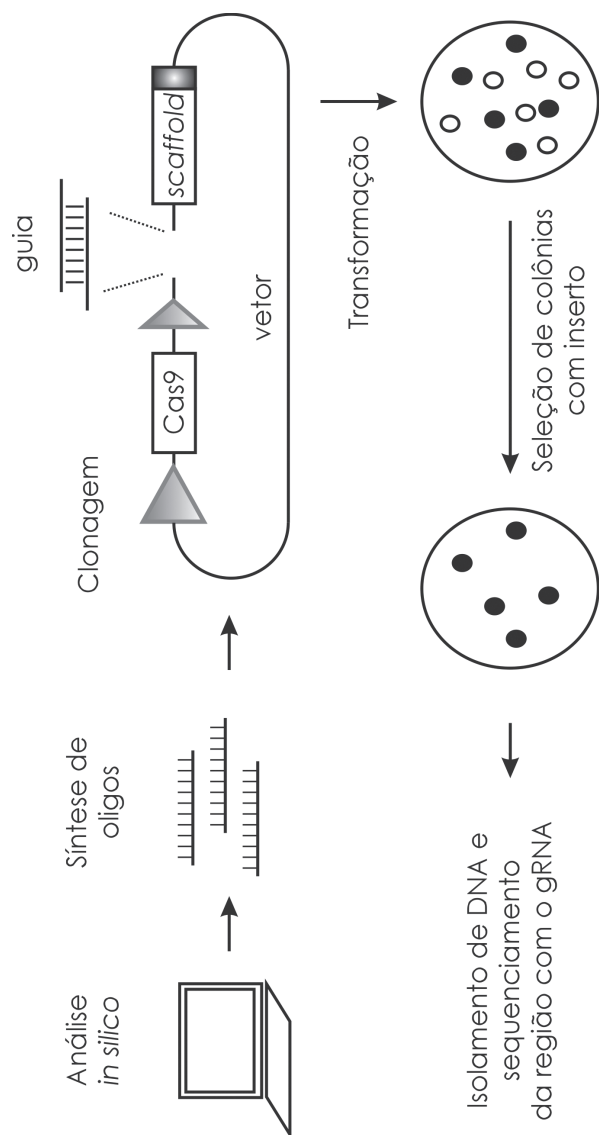


Figura 4. Visão geral dos passos a serem seguidos para a construção de bibliotecas de gRNAs. Após a obtenção das sequências-guias *in silico* e a síntese dos oligonucleotídeos correspondentes, estas moléculas devem ser integradas ao vetor (por digestão/ligação ou LIG). Em seguida, o vetor deve ser inserido em bactérias competentes via transformação, purificado e sequenciado. Eventuais etapas posteriores, como a produção de partículas (lenti)virais, podem ser aplicadas. Imagem: Gonçalves, NJN e Pereira, TC (2016).

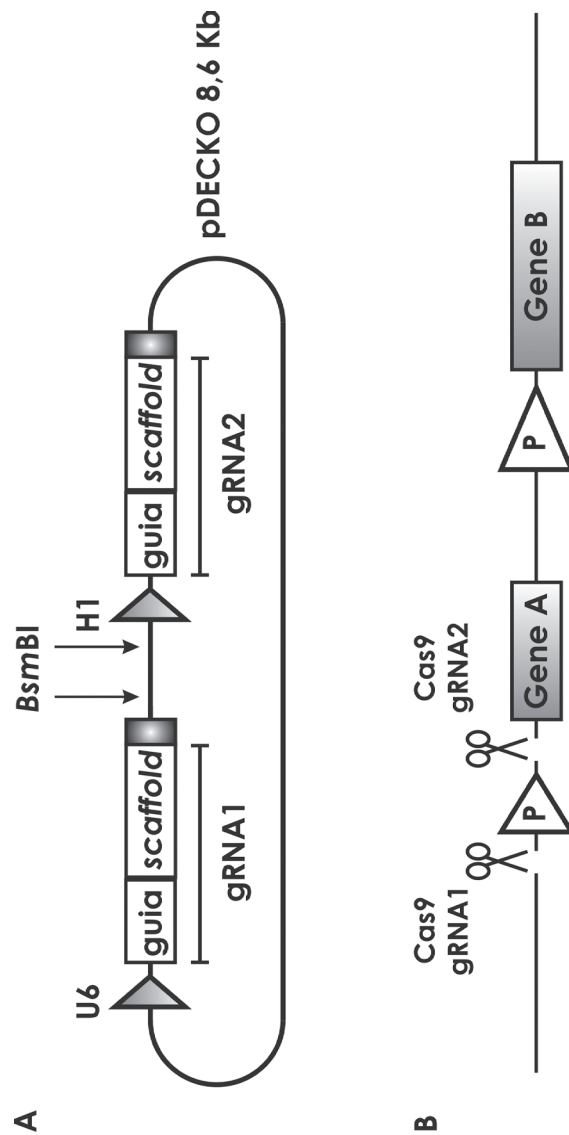


Figura 5. DECKO. A) Estrutura do plasmídeo DECKO, com respectivos sítios de clonagem (BsmBI), no qual dois gRNAs independentes (promotores distintos - U6 e H1) são direcionados de acordo com suas sequências-guia. B) Deleção de um elemento genético (promotor do gene A), mediado por nucleases guiadas pelos gRNAs1 e 2. P: promotor. Imagem: Gonçalves, NJN e Pereira, TC (2016).

6.2 Tipos de *pooled libraries*

Bibliotecas de pares de sgRNAs com promotores distintos

Um crescente número de aplicações da CRISPR requer a expressão simultânea de dois sgRNAs, transcritos por promotores diferentes. Uma vez que as sequências devem ser clonadas em regiões promotoras diferentes no mesmo plasmídeo, as estratégias já disponíveis e validadas para construção de bibliotecas não são efetivas neste caso. Vidigal e Ventura (2015) desenvolveram um método rápido e eficiente para obtenção de bibliotecas de pares. Resumidamente, as duas sequências foram clonadas simultaneamente no vetor, a partir de um oligonucleotídeo de 110 nt. Este desenho contém a sequência dos dois gRNAs separadas por sítios de ligação da enzima *BbsI*, e nas pontas 5' e 3' opostas existem sequências curtas de homologia com o promotor U6 e com a *sequência universal* do gRNA (o *scaffold*). Em paralelo, foi gerado um plasmídeo doador que, digerido com *BbsI*, possibilita a ligação e posterior clonagem das sequências gRNAs.

Um belo exemplo de gRNAs pareados refere-se à metodologia DECKO (*Double Excision CRISPR Knockout*). Esta é uma estratégia eficaz em que vetores lentivirais expressam simultaneamente dois gRNAs para a deleção completa de algum elemento genético de localização conhecida (promotor, *enhancer*, região codificadora, etc.) (figura 5). DECKO se baseia em um oligonucleotídeo inicial de 165 pb, carreando as sequências variáveis (*guias*) dos dois gRNAs, permitindo assim uma dupla excisão e, conseqüentemente, a deleção completa de uma região genômica de interesse (Aparicio-Prat et al., 2015). Esta simplicidade permite que DECKO seja escalonável para uso como biblioteca.

Bibliotecas para análise de elementos cis-regulatórios (Molecular Chipper)

Cheng et al. (2016) obtiveram uma bibliotecas de sgRNAs de alta cobertura (*i.e.*, muitos alvos por kilobase) por meio da abordagem *molecular chipper*, que consiste em uma associação de fragmentação aleatória do DNA em estudo e uso da enzima de restrição *EcoP15I* para posterior clonagem. Esta estratégia é interessante quando não se sabe a localização precisa do(s) elemento(s) genético(s) em estudo. Um exemplo clássico é a análise da região regulatória *upstream* de um gene, que pode ser composta por diversos pequenos domínios (ou *motivos*) espalhados ao longo de várias centenas (ou milhares) de pares de bases. Nesse sentido, é necessário gerar diversos gRNAs para esta região e testá-los individualmente quanto ao seu efeito na expressão gênica.

Uma biblioteca construída via *molecular chipper* foi utilizada para identificar novos domínios regulatórios de um grupo de 17 microRNAs, uma importante classe de genes não codificadores que regulam diversos processos biológicos. O miR-142, por exemplo, tem um papel muito importante na biogênese de células hematopoiéticas e mutações nesta região foram associadas à formação tumoral. A estratégia de *molecular chipper* permitiu a obtenção de um conjunto de sgRNAs cobrindo toda a potencial região regulatória do miR-142 3p, para o escaneamento por elementos regulatórios da transcrição (*cis-acting elements*). Tais sgRNAs promoveram mutações pontuais ao longo de toda esta região, resultando em alterações parciais ou completas na expressão gênica, avaliada por meio da atividade do repórter GFP e FACS (figura 6). Assim, foi possível um escaneamento de elementos regulatórios da transcrição conhecidos e não conhecidos de miR-142.

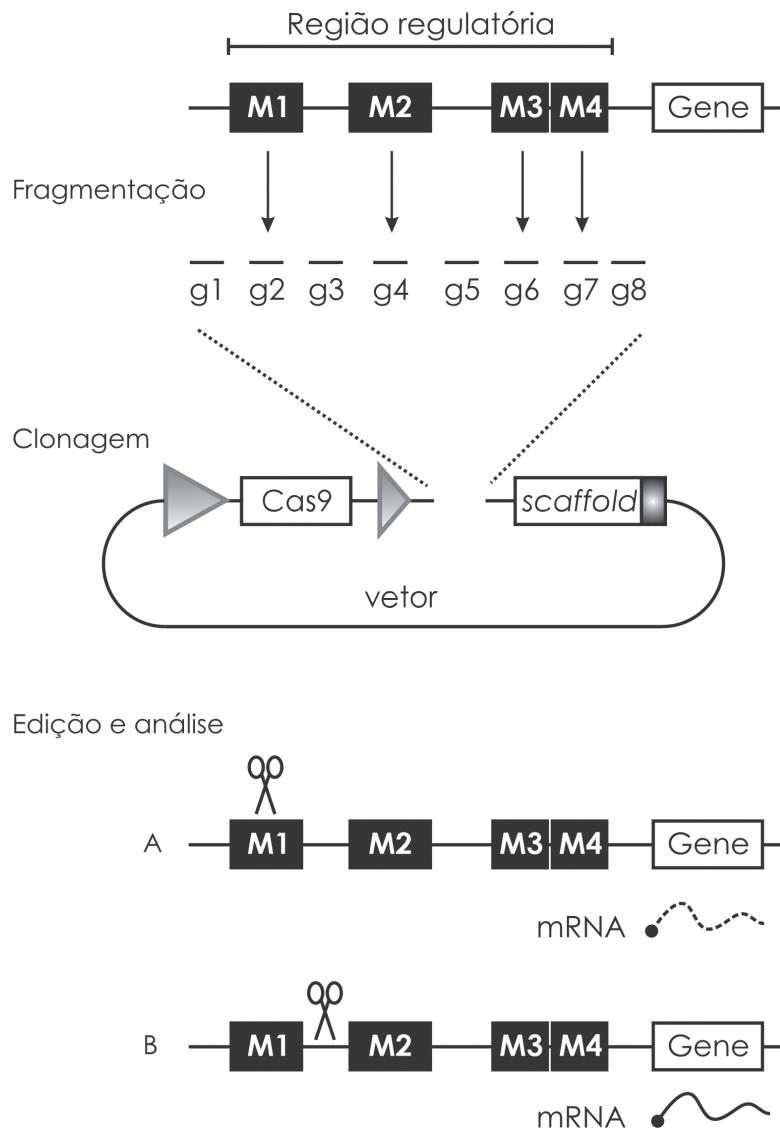


Figura 6. Princípio da estratégia Molecular Chipper. Esta abordagem visa gerar gRNAs referentes a determinada região do genoma na qual se sabe existir elementos genéticos (como motivos funcionais – M1 a M4 – dentro de uma região regulatória de um gene), mas não se sabe suas localizações precisas. Os diversos fragmentos gerados (g1 a g8) podem ser clonados no vetor CRISPR, gerando uma biblioteca. Alguns dos clones (g2, g4, g6 e g7) são derivados diretamente dos motivos funcionais e resultarão em edições nos mesmos. Isto, por sua vez, alterará a expressão gênica, avaliada por um gene repórter (A). Note que g3 (B) não atinge um motivo funcional, não afetando assim a expressão gênica. Imagem: Gonçalves NJN e Pereira TC (2016).

Bibliotecas para CRISPRi e CRISPRa

É possível também construir bibliotecas para alterar a expressão de genes via CRISPR. Nesse caso, os gRNAs devem ser clonados em vetores contendo uma Cas9 especial, capaz de reprimir ou ativar a transcrição gênica (vide cap. 3). Para estudos em células eucarióticas, por exemplo, a dCas9 pode ser fusionada com o domínio de ativação de p65 ou VP64. Estudos têm demonstrado que dCas9-VP64 é capaz de ativar genes endógenos silenciados ou intensificar a expressão de genes já ativos (Russa e Qi, 2015). Tal indução transcricional pode ser aprimorada usando múltiplos sgRNAs contra o mesmo alvo.

Gilbert e colaboradores (2014) investigaram a repressão (CRISPRi, usando 206.421 sgRNAs) ou ativação (CRISPRa, usando 198.810 gRNAs) gênicas, e os resultados mostram controle robusto com níveis baixos de *off-targets* em células humanas K562, avaliando 15.977 genes (Gilbert et al., 2014).

Outra abordagem utiliza o sistema SAM (*synergistic activation mediator*), que implementa múltiplos ativadores transcricionais para criar um efeito sinérgico. Essa ferramenta faz uso do dCas9-VP64, porém os autores arquitetaram características adicionais para reforçar o recrutamento de ativadores. Nesta nova abordagem, o sgRNA contém aptâmeros que permitem a ligação da proteína MS2 (de bacteriófago) fusionada a dois outros domínios de ativação (p65 e HSF1). O sistema SAM tem um alcance maior e mais aperfeiçoado para a ativação transcricional e com ele foram avaliados 23.430 genes com 70.290 sgRNAs em células A375 (melanoma) (Konermann et al., 2015).

Bibliotecas codificadas (barcoded)

Uma nova abordagem foi descrita por pesquisadores que geraram uma plataforma associativa da análise combinatória em massa (CombiGEM) com CRISPR, para uma montagem rápida e precisa de bibliotecas que podem ser mapeadas com sequenciamento em larga escala. Esta abordagem pode ser muito útil para identificar perturbações em células humanas e para identificar grupos de genes que controlam características complexas (Kock et al., 2016). Wong et al. (2016) usaram essa abordagem para avaliar uma grande combinação de genes em células humanas, traduzindo esses dados em uma combinação efetiva de drogas específicas. A partir dessa metodologia foi gerada uma biblioteca codificada (*barcodes*) de 23.409 sequências duplas de gRNA, e então usou-se sequenciamento em larga escala para avaliar quais genes-alvos foram capazes de inibir o crescimento de células cancerosas no ovário. Tal efeito foi validado para um conjunto específico de genes que incluem reguladores epigenéticos. A combinação de tais tecnologias para a obtenção de bibliotecas vai acelerar de forma sistemática a identificação e modulação de drogas específicas para diversas doenças humanas de alta complexidade.

7. Conclusões

O rápido avanço da tecnologia CRISPR se traduz no oferecimento de uma extensa lista de produtos acadêmica e comercialmente disponíveis, assim como o desenvolvimento de ferramentas *in silico* e estratégias moleculares para a síntese e/ou construção dos *reagentes de CRISPR* (Cas9 e gRNA). Os recursos para obtenção dessas moléculas é vasto, podendo ser ajustado de acordo com os objetivos específicos do pesquisador.

Referências

- Ablain J, Durand EM, Yang S, Zhou Y, Zon LIA. CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev Cell*. 2015;32(6):756-64.
- Agrotis A, Ketteler R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front Genet*. 2015;6(300):1-15.
- Aida T, Chiyo K, Usami T, Ishikubo H, Imahashi R, Wada Y, et al. Cloning-free CRISPR/Cas system facilitates functional cassette knock-in in mice. *Genome Biol*. 2015;16(1):87.
- Ali Z, Abul-Faraj A, Li L, Ghosh N, Piatek M, Mahjoub A, et al. Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. *Mol Plant*. 2015;8(8):1288-91.
- Ansai S, Kinoshita M. Targeted mutagenesis using CRISPR/Cas system in medaka. *Biol Open*. 2014;3(5):362-71.
- Aparicio-Prat E, Arnan C, Sala I, Bosch N, Guigó R, Johnson R. DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. *BMC Genomics*. 2015;16(846):1-15.
- Bae S, Kweon J, Kim HS, Kim JS. Microhomology-based choice of Cas9 nuclease target sites. *Nat Methods*. 2014 Jul;11(7):705-6.
- Bassett AR, Kong L, Liu JL. A Genome-Wide CRISPR Library for High-Throughput Genetic Screening in *Drosophila* Cells. *J Genet Genomics*. 2015;42(6):301-9.
- Bhattacharya D, Marfo CA, Li D, Lane M, Khokha MK. CRISPR/Cas9: an inexpensive, efficient loss of function tool to screen human disease genes in *Xenopus*. *Dev Biol*. 2015;408(2):196-204.
- Bhaya D, Davison M, Barrangou R. CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annu Rev Genet*. 2011;45:273-97.
- Blin K, Pedersen LE, Weber T, Lee SY. CRISPy-web: An online resource to design sgRNAs for CRISPR applications. *Synthetic and Systems Biotechnology* 2016.
- Blitz IL, Biesinger J, Xie X, Cho KW. Biallelic genome modification in F0 *Xenopus tropicalis* embryos using the CRISPR/Cas system. *Genesis*. 2013;51(12):827-34.
- Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*. 2013;23(10):1163-71.
- Cheng J, Roden CA, Pan W, Zhu S, Baccei A, Pan X, et al. A Molecular Chipper technology for CRISPR sgRNA library generation and functional mapping of noncoding regions. *Nat Commun*. 2016;7(11178):1-10.
- Cho SW, Kim S, Kim JM, Kim J. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*. 2013;31(3):230-2.
- Chu VT, Weber T, Graf R, Sommermann T, Petsch K, Sack U, et al. Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. *BMC Biotechnol*. 2016;16(1):1.
- Cong L, Ran A, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*. 2013;339:819-23.
- Dalgin G, Prince VE. Differential levels of Neurod establish zebrafish endocrine pancreas cell fates. *Dev Biol*. 2015;402(1):81-97.
- Deng W, Shi X, Tjian R, Lionnet T, Singer RH. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc Natl Acad Sci USA*. 2015;112(38):11870-5.
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*. 2016;34:184-91.
- Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol*. 2014 Dec;32(12):1262-7.
- Fu Y, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013;31:822-6.
- Fujiwara Y, Ikawa M. CRISPR/Cas9-Based Genome Editing in Mice by Single Plasmid Injection. *Methods Enzymol*. 2014;546:319-36.
- Gagnon JA, Valen E, Thyme SB, Huang P, Ahkmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One*. 2014;9(5):e98186.
- Gao Y, Zhang Y, Zhang D, Dai X, Estelle M, Zhao Y. Auxin binding protein 1 (ABP1) is not required for either auxin signaling or Arabidopsis development. *Proc Natl Acad Sci USA*. 2015 Feb;112(7):2275-80.
- Garneau JE, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010;468:67-71.
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014;159:647-61.
- Guo X, Zhang T, Hu Z, Zhang Y, Shi Z, Wang Q, et al. Efficient RNA/Cas9-mediated genome editing in *Xenopus tropicalis*. *Development*. 2014;141(3):707-14.
- Heigwer F, Zhan T, Breinig M, Winter J, Brügemann D, Leible S, et al. CRISPR library designer (CLD): software for multispecies design of single guide RNA libraries. *Genome Biol*. 2016;17(55):1-10.
- Hisano Y, Sakuma T, Nakade S, Ohga R, Ota S, Okamoto H, et al. Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep*. 2015:5.
- Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA*. 2013 Sep;110(39):15644-9.
- Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA*. 2013 Sep;110(39):15644-9.
- Hruscha A, Krawitz P, Reichenberg A, Heinrich V, Hecht J, Haass C, et al. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*. 2013;140(24):4982-7.
- Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827-32.
- Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res*. 2013:gkt780.

- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012;337:816-21.
- Jr VAB, Zarecor S, Wright DA, Wang Y, Liu J, Chen K, et al. A quick guide to CRISPR sgRNA design tools. *GM Crops Food*. 2016;6:266-76.
- Kalebic N, Taverna E, Tavano S, Wong FK, Suchold D, Winkler S, et al. CRISPR/Cas9-induced disruption of gene expression in mouse embryonic brain and single neural stem cells in vivo. *EMBO Rep*. 2016:e201541715.
- Kampmann M, Bassik MC, Weissman JS. Functional genomics platform for pooled screening and generation of mammalian genetic interaction maps. *Nat Protoc*. 2014;9(8):1825-47.
- Kiani S, Chavez A, Tuttle M, Hall RN, Chari R, Ter-Ovanesyan D, et al. Cas9 gRNA engineering for genome editing, activation and repression. *Nat Methods*. 2015;12(11):1051-4.
- Kim H, Kim S, Ryu J, Choi MK, Kweon J, Kang B, et al. A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. *J Integr Plant Biol*. 2016. [Epub ahead of print]
- Koch L. CombiGEM-CRISPR: a creative combination. *Nat Rev Genet*. 2016;17(194)
- Koike-Yusa H, Li Y, Tan E, VelascoHerrera MDC, Yusa K. Genomewide recessive genetic screening in mammalian cells with a lentiviral CRISPRguide RNA library. *Nat Biotechnol*. 2013;32:267-73.
- Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015;517:583-8.
- Kotani H, Taimatsu K, Ohga R, Ota S, Kawahara A. Efficient multiple genome modifications induced by the crRNAs, tracrRNA and Cas9 protein complex in zebrafish. *PLoS One*. 2015;10(5):e0128319.
- Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genomewide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014 Jul;32(7):677-83.
- Lane AB, Strzelecka M, Ettinger A, Grenfell AW, Wittmann T, Heald R. Enzymatically Generated CRISPR Libraries for Genome Labeling and Screening. *Dev Cell*. 2015;34:373-8.
- Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*. 2013 Nov;8(11):2180-96.
- Ledford H. Crispr, the Disruptor. *Nature*. 2015;522:20-4.
- Lee CM, Cradick TJ, Fine EJ, Bao G. Nuclease Target Site Selection for Maximizing On-target Activity and Minimizing Off-target Effects in Genome Editing. *Mol Ther*. 2016 Mar;24(3):475-87.
- Lee NC, Larionov V, Kouprina N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res*. 2015;43(8):e55-55.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823-6.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. *Science*. 2013;339:823-6.
- Malina A, Katigbak A, Cencic R, Maiga RI, Robert F, Miura H, et al. Adapting CRISPR/Cas9 for Functional Genomics Screens. *Methods Enzymol*. 2014;546:193-213.
- Mao Y, Zhang Z, Feng Z, Wei P, Zhang H, Botella JR, et al. Development of germ-line-specific CRISPR/Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. *Plant Biotechnol J*. 2015.
- Mohr SE, Hu Y, Ewen-Campen B, Housden BE, Viswanatha R, Perrimon N. CRISPR guide RNA design for research applications. *FEBS J*. 2016 Jun
- Park CY, Kim DH, Son JS, Sung JJ, Lee J, Bae S, et al. Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. *Cell Stem Cell*. 2015;17(2):213-20.
- Pattanayak V, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*. 2013;31:839-43.
- Peng J, Zhou Y, Zhu S, Wei W. High-throughput screens in mammalian cells using the CRISPR-Cas9 system. *FEBS J*. 2015;282:2089-96.
- Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell*. 2014 Oct;159(2):440-55.
- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015 Apr;520(7546):186-91.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-308.
- Ren Q, Li C, Yuan P, Cai C, Zhang L, Luo GG, et al. A Dual-Reporter System for Real-Time Monitoring and High-throughput CRISPR/Cas9 Library Screening of the Hepatitis C Virus. *Sci Rep*. 2015;5(8865):1-6.
- Russa MFL, Qi LS. The new state of the art: CRISPR for gene activation and repression. *Mol Cell Biol*. 2015;36(9):1-42.
- Ryan OW, Cate JHD. Multiplex Engineering of Industrial Yeast Genomes Using CRISPRm. *Methods Enzymol*. 2014;546:473-89.
- Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, Lee ME, DeLoache W, Dueber JE, Arkin AP, Cate JHD. Selection of chromosomal DNA libraries using a multiplex CRISPR system. 2014. 1-15
- Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11(8):783-4.
- Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11(8):783-4.
- Schmid-Burgk JL, Schmidt T, Kaiser V, Höning K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol*. 2013;31(1):76-81.
- Schmidt T, Schmid-Burgk JL, Hornung V. Synthesis of an arrayed sgRNA library targeting the human genome. *Sci Rep*. 2015;5(14987):1-10.
- Schwartz ML, Jorgensen EM. SapTrap, a Toolkit for High-Throughput CRISPR/Cas9 Gene Modification in *C. elegans*. *Genetics*. 2016;202(4):1277-88.
- Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. *Natl Rev*. 2015;16:299-311.

- Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol*. 2013;31(8):686-8.
- Sheridan RM, Bentley DL. Selectable one-step PCR-mediated integration of a degron for rapid depletion of endogenous human proteins. *Biotechniques*. 2016;60:69-74.
- Shi Z, Wang F, Cui Y, Liu Z, Guo X, Zhang Y, et al. Heritable CRISPR/Cas9-mediated targeted integration in *Xenopus tropicalis*. *FASEB J*. 2015;29(12):4914-23.
- Smith JD, Suresh S, Schlecht U, Wu M, Wagih O, Peltz G, et al. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol*. 2016;17(45):1-16.
- Stone EA. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11(8):783-4.
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159(3):635-46.
- Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*. 2015 Feb;33(2):187-97.
- Varshney GK, Pei W, LaFave MC, Idol J, Xu L, Gallardo V, et al. High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res*. 2015;25(7):1030-42.
- Vidigal JA, Ventura A. Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. *Nat Commun*. 2015;6(8083):1-7.
- Wang T, Lander ES, Sabatini DM. Large-Scale Single-Guide RNA Library Construction and Use for Genetic Screens: CRISPR/Cas9-based Genetic Screens. *Cold Spring Harb Protoc*. 2016. [Epub ahead of print].
- Wang T, Lander ES, Sabatini DM. Protocol 1: Single-Guide Library Design and Construction. *Cold Spring Harb Protoc*. 2016. [Epub ahead of print].
- Wang T, Lander ES, Sabatini DM. Protocol 2: Viral Packaging and Cell Culture for CRISPR-based Screens. *Cold Spring Harb Protoc*. 2016;2016(4):1-8.
- Wang X, Zhou J, Cao C, Huang J, Hai T, Wang Y, et al. Efficient CRISPR/Cas9-mediated biallelic gene disruption and site-specific knockin after rapid selection of highly active sgRNAs in pigs. *Sci Rep*. 2015;5.
- Wong ASL, Choia GCG, Cuia CH, Pregernig G, Milanid P, Adam M, et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombGEM. *Proc Natl Acad Sci USA*. 2016;113(9):2544-9.
- Wong N, Liu W, Wang X. WU-CRISPR: characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biol*. 2015;16(218):1-8.
- Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*. 2014 Jul;32(7):670-6.
- Yang Y, Wang L, Bell P, McMenamin D, He Z, White J, et al. A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat Biotechnol*. 2016 Mar;34(3):334-8.
- Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*. 2015;160(1):339-50.
- Zhang H, Zhang X, Fan C, Xie Q, Xu C, Zhao Q, et al. A novel sgRNA selection system for CRISPR-Cas9 in mammalian cells. *Biochem Biophys Res Commun*. 2016;471:528-32.
- Zhang Z, Mao Y, Ha S, Liu W, Botella JR, Zhu JK. A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in *Arabidopsis*. *Plant Cell Rep*. 2015;1-15.
- Zhong C, Yin Q, Xie Z, Bai M, Dong R, Tang W, et al. CRISPR-Cas9-Mediated Genetic Screening in Mice with Haploid Embryonic Stem Cells Carrying a Guide RNA Library. *Cell Stem Cell*. 2015;17(2):1-12.

Controles experimentais, formas de análise e cuidados especiais

Capítulo 8

Bruno Trevenzoli Favero¹, Oliver Kemp¹ e Tiago Campos Pereira^{2,3}

¹*Dept. of Plant and Environmental Sciences, Universidade de Copenhagen - Dinamarca*

²*Dept.º de Biologia, FFCLRP, USP, Ribeirão Preto, SP - Brasil*

³*Programa de Pós-Graduação em Genética, FMRP - USP, Ribeirão Preto, SP - Brasil*

Estruturação do capítulo

1. Introdução
2. Controles experimentais
 - 2.1 Controles negativos
 - 2.2 Controles positivos
 - 2.3 De especificidade
3. Confirmação da edição
 - 3.1 Análises de DNA
 - 3.2 Análises de RNA
 - 3.3 Análise da Proteína
4. Cuidados especiais
 - 4.1 Técnicas para identificação de *off-targets*
 - 4.2 Efeito mosaico
5. Conclusões

1. Introdução

Controles experimentais são importantes para a interpretação correta dos resultados de qualquer estudo. Apesar da grande versatilidade da técnica (vide cap. 3), neste capítulo iremos focar nos principais controles e formas de análise empregados durante a *edição genética* mediada por CRISPR, que é a aplicação mais frequente desta.

Na edição genômica de precisão, a hipótese é a de que uma modificação de um local específico no genoma do organismo em estudo é a responsável pela alteração fenotípica observada pós-experimento. Os controles e análises moleculares abordados neste capítulo visam garantir que as conclusões derivadas dos estudos que utilizam CRISPR estejam precisas.

2. Controles experimentais

2.1 Controles negativos

A forma de controle negativo mais utilizada é a *administração incompleta* dos elementos necessários para a edição. Ou seja, Cas9 e/ou o gRNA são excluídos do procedimento (*e.g.*, vetor vazio) para mostrar que os eventos de quebra de fita dupla são dependentes de ambas as partes do sistema. Alternativamente, o uso de Cas9 e um gRNA contra uma sequência não presente no genoma também é comumente utilizado como controle negativo (às vezes denominado “*universal negative control*”). Este serve para evidenciar que o processo é sequência-específico, ou seja, apenas o gRNA direcionado ao seu gene é capaz de inativar o alvo.

Já para os ensaios de KI e/ou substituição alélica, envolvendo um DNA doador exógeno e HDR, o controle negativo tipicamente é caracterizado pelo mesmo procedimento conduzido na ausência do DNA doador.

2.2 Controles positivos

O uso de um controle positivo, *i.e.*, um sistema simples cuja funcionalidade já está comprovada (*e.g.*, um plasmídeo com Cas9 e gRNA validados na espécie de estudo) pode ser interessante em alguns casos. Assim, pesquisadores que utilizam a técnica pela primeira vez e que desejam “testá-la e/ou excluir possíveis fontes de erro ao longo do procedimento”, podem alcançar um resultado positivo garantido. Experimentos-piloto podem então ser executados para confirmar que todas as outras condições laboratoriais estão adequadas para se obter a edição dos genes mediada por CRISPR. No entanto, isto requer que os componentes do piloto sejam os mesmos que os utilizados no estudo, por exemplo, vetores, promotores, tipo de célula, etc.

Como alternativa, em vez de adquirir comercialmente um “controle positivo”, pode-se construir em laboratório um sistema com um gRNA previamente validado em estudos publicados (e cuja sequência está disponível em artigos).

2.3 De especificidade

Como mencionado anteriormente, problemas de especificidade ocorrem mesmo ao se usar ferramentas de edição genômica de precisão, tais como CRISPR. A perda de especificidade deve ser entendida como eventos de DSBs ocorridos em sítio(s) que não era(m)

o alvo pretendido (*off-targets*). Acredita-se que isto possa ser devido a semelhanças entre a sequência-guia e outras porções no genoma. A gravidade da perda de especificidade é naturalmente dependente da finalidade e objetivos da pesquisa. Em estudos de identificação da função gênica, o nocauteamento de outros genes pode comprometer a interpretação dos resultados. Na terapia gênica de células humanas via CRISPR, edição de *off-targets* não são um risco aceitável.

A especificidade de um gRNA pode ser acompanhada em pelo menos quatro níveis diferentes. O primeiro deles diz respeito às análises *in silico*, ainda durante a etapa de desenho do gRNA, à procura de regiões no genoma semelhantes à sequência-guia, *i.e.*, potenciais *off-targets* (cap. 7). O segundo se refere a controles de especificidade *in si* (vistos a seguir). O terceiro se refere a técnicas que permitem avaliar, pós-experimento de edição, eventuais inespecificidades de Cas9/gRNA (seção 4.1 deste capítulo). Por fim, estudos de complementação com um alelo funcional restaurando o fenótipo selvagem podem confirmar o papel do gene de Interesse (GDI) (Sander e Joung, 2014).

A inclusão de controles de especificidade é dependente dos detalhes de cada desenho experimental. Consideremos um caso simples, em que se objetiva nocautear determinado GDI e observar a alteração fenotípica resultante. Neste caso, o nocauteamento acidental e simultâneo de um eventual *off-target* pode interferir no fenótipo, comprometendo a interpretação do experimento.

Tabela 1. Controles experimentais mais comuns em estudos de edição genética utilizando CRISPR.

Controles experimentais	Edição genética			
	<i>knockout</i>	<i>knock-in</i>	substituição alélica	deleção gênica
Controles negativos	Administração incompleta do sistema (sem Cas9 ou sem gRNA)	Administração incompleta do sistema (tipicamente sem o DNA doador)	Administração incompleta do sistema (tipicamente sem o DNA doador)	Administração incompleta do sistema (sem Cas9 ou sem gRNA)
	Cas9 e gRNA contra sequência inexistente no genoma (<i>e.g.</i> , <i>universal negative control</i>)	–	–	–
Controle positivo	Cas9 e gRNA já validado experimentalmente (<i>e.g.</i> , <i>EMX1 Positive Control</i>)	Não é comum	Não é comum	Não é comum
Controles de especificidade	Dois ou mais gRNAs contra sítios diferentes do mesmo gene-alvo	Não é comum	Não é comum	Não é comum

Nota. Os controles devem ser adaptados de acordo com a estratégia de administração utilizada, *i.e.*, mesmo tipo de molécula (vetor, mRNA ou proteína) e mesmo tipo de Cas9 (Cas9, Cas9n, dCas9, etc.).

“Controles de especificidade” dizem respeito a um conjunto de dois ou mais outros gRNAs direcionados contra diferentes regiões do mesmo GDI. Considere que o gRNA-1, por exemplo, promova duas alterações fenotípicas: redução da taxa de proliferação celular e redução do tamanho celular. O gRNA-2: sensibilidade a altas temperaturas e redução da taxa de proliferação celular. O gRNA-3: redução da taxa de proliferação celular e alterações na morfologia do envelope nuclear. Conclui-se, desta forma, que o nocauteamento do GDI resulta na redução da taxa de proliferação celular, pois esta é a alteração fenotípica comum a todos os três gRNAs. As outras alterações são devidas *ao nocauteamento de diferentes off-targets* (cada gRNA apresenta *off-targets* distintos). Uma síntese dos controles experimentais e suas aplicações na edição gênica podem ser vistos na Tabela 1.

3. Confirmação da edição

Um experimento bem-sucedido de CRISPR irá promover uma alteração no DNA. Dependendo da natureza desta edição genética, ela *poderá resultar em alterações qualitativas e/ou quantitativas no RNA e/ou proteína em maior ou menor extensão*. É crucial compreender que diferentes abordagens (KO/KI/SA/DG) ou até mesmo o local de mutação dentro do gene (vide cap. 7, figura 2) podem promover efeitos distintos nas três ‘camadas’ (DNA, RNA e proteínas). Apresentaremos a seguir algumas técnicas que permitem a confirmação da edição genética considerando estes três níveis. Uma síntese destas técnicas pode ser vista na Tabela 2.

Tabela 2. Técnicas para confirmação molecular da edição.

Molécula analisada	Nível de análise	Técnica	Edição genética			
			<i>knockout</i>	<i>knock-in</i>	substituição alélica	deleção gênica
DNA	Global	Sequenciamento em larga escala	Sim	Sim	Sim	Sim
		PCR + EMC	Sim	NA	Sim	Não é comum
		PCR + EDSR	Sim	NA	Não é comum	Não é comum
		IDAA	Sim	NA	Não é comum	Não é comum
	Triagem		Apenas se a edição resultar em alterações dramáticas no tamanho do gene	Amplificação do transgene	Amplificação usando sondas alelo-específicas	Confirmação da ausência do gene
		PCR				

NA: não se aplica.

EDSR: Ensaio de interrupção de sítio de estrição.

Tabela 2. Continuação.

Molécula analisada	Nível de análise	Técnica	Edição genética			
			<i>knockout</i>	<i>knock-in</i>	substituição alélica	deleção gênica
Proteína	Específico	<i>western blot</i>	Possível se houver alteração no tamanho ou abundância da proteína	Detecção da proteína heteróloga	Possível se houver alteração no tamanho ou abundância da proteína	Confirmação da ausência da proteína
		ELISA	Possível se houver alteração na abundância da proteína	Detecção da proteína heteróloga	Possível se houver alteração na abundância da proteína	Confirmação da ausência da proteína
		Ensaio bioquímico	Possível se houver perda da atividade bioquímica	Detecção da proteína heteróloga	Possível se houver perda da atividade bioquímica	Confirmação da ausência da proteína

NA: não se aplica.

EDSR: Ensaio de disrupção de sítio de estrição.

3.1 Análises de DNA

A fim de determinar se as mudanças no sítio-alvo ocorreram, o sequenciamento em larga escala é comumente aplicado, já que muitas vezes o evento de edição é raro ou diluído na população de células ou tecidos do organismo. Entretanto, o sequenciamento em larga escala pode não ser uma opção economicamente viável.

Nesse sentido, técnicas baratas e eficientes de triagem podem ser úteis para analisar o DNA. As amostras selecionadas após esta etapa de triagem podem então ser sequenciadas e gerar uma análise mais precisa e detalhada. A seguir são apresentadas algumas técnicas para triagem pré-sequenciamento de amostras promissoras.

EMC

O ensaio enzimático de clivagem de bases não pareadas (*Enzymatic Mismatch Cleavage* – EMC) tem sido utilizado para triagem de mutações, independentes de sua natureza (deleções, inserções ou substituições) (figura 1). Estes ensaios são mencionados nos artigos científicos de diversas formas, mas o conceito é o mesmo. Inicialmente amplifica-se por PCR a região correspondente ao sítio-alvo, tanto em amostras experimentais (potencialmente editadas) quanto em amostras do controle negativo (não editadas). Ao se misturar ambos os amplicons, caso a edição tenha ocorrido, serão formados *heteroduplexes* (*i.e.*, uma cadeia mutada pareando-se com uma cadeia não mutada, resultando em mal pareamentos). As deleções/inserções vão causar pequenas protuberâncias devido às bases não pareadas no heteroduplex, que são então digeridos com enzimas que reconhecem essas inconsistências e cortam a sequência. Em seguida, o DNA será separado por eletroforese em gel, resultando em múltiplas bandas se a

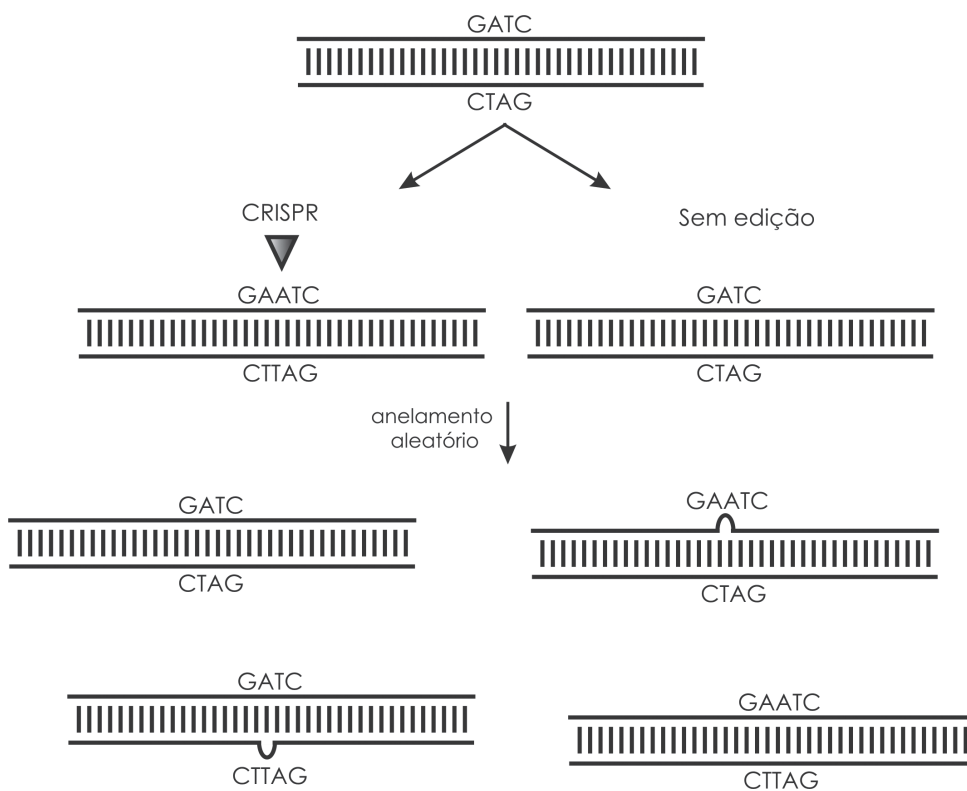


Figura 1. Princípio do EMC (ensaio enzimático de clivagem de bases não pareadas). Nesta estratégia, o DNA de células submetidas a CRISPR (portanto, possivelmente editado) é misturado com DNA de amostras não submetidas a CRISPR (não editadas). Após o aquecimento e resfriamento lento, as cadeias de DNA irão hibridizar aleatoriamente, gerando eventuais duplexes com bases não pareadas (indicados com as pequenas alças). Estes *mismatches* serão clivados enzimaticamente (por CEL, ENDO ou T7E1), gerando produtos visíveis em gel. Triângulo invertido: clivagem mediada por Cas9. Imagem: Pereira, TC (2016).

dupla fita de DNA apresentava protuberâncias. As nucleases CEL, ENDO e T7E1 (*e.g.*, *Surveyor Mutation Detection Kit*) são exemplos de enzimas que têm sido utilizadas para esses tipos de ensaios, que são de baixo custo. Mas tem havido um debate sobre a sensibilidade com que esta metodologia detecta as edições (Vouillot et al., 2015).

Ensaio de interrupção de sítio de restrição

Outro método amplamente utilizado se baseia em direcionar Cas9 a um sítio de uma enzima de restrição (ER) localizado no DNA-alvo (figura 2) (Hruscha e Schmid, 2015). Caso a edição tenha sido bem-sucedida, a mutação irá perturbar o sítio de reconhecimento da ER e o respectivo amplicon (gerado via PCR) será resistente à digestão por esta ER (analisado em gel de agarose). Este método é sem dúvida mais sensível do que o EMC, mas requer que CRISPR seja concebida para atingir um sítio ER dentro do gene-alvo. Este parâmetro adicional no desenho da *sequência-guia* pode eventualmente ser um complicador.

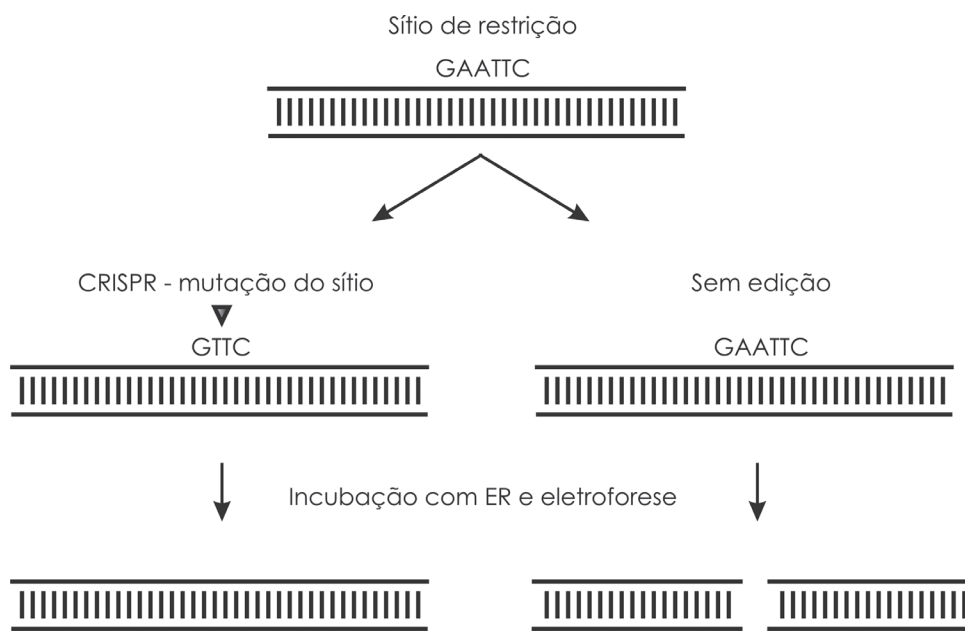


Figura 2. Princípio do ensaio de disruptão de sítio de restrição. Nesta estratégia, o gRNA é desenhado para editar um sítio de reconhecimento de uma endonuclease de restrição, localizado dentro do gene-alvo. Após o experimento, o gene-alvo é amplificado por PCR e submetido a digestão com a respectiva enzima de restrição. Caso a edição tenha sido bem-sucedida, o sítio de reconhecimento terá sido alterado, não gerando produtos de digestão visíveis em gel. ER: enzima de restrição. Triângulo invertido: clivagem mediada por Cas9. Imagem: Pereira, TC (2016).

IDAA

A detecção de *indels* por análise de amplicon (*Indel Detection by Amplicon Analysis* – IDAA) é um método ligeiramente mais sofisticado. Neste, a região do DNA-alvo editado é amplificada por PCR, sendo que o *primer* direto contém uma extensão que complementa um *primer* universal marcado com fluoresceína. Isto permite a marcação de amplicons com fluoresceína e a separação por tamanho em eletroforese capilar. Esse método permite não apenas a detecção de *indels* como também permite a caracterização do tamanho destas, sejam deleções ou inserções. Contudo, isso exige um equipamento mais caro do que os métodos mencionados anteriormente (Yang et al., 2015).

3.2 Análises de RNA

A maioria dos estudos de edição genética envolvem pequenas modificações nucleotídicas que não alteram drasticamente as características do mRNA (tamanho ou abundância). Dessa forma, as análises convencionais de RT-(q)PCR são dispensáveis na maioria desses estudos.

Contudo, elas podem ter valia em pesquisas envolvendo: (i) manipulação de sequências regulatórias, (ii) CRISPRa ou CRISPRi, (iii) clivagem de RNA, (iv) modificação em sítios de *splicing*, (v) grandes alterações estruturais no gene (deleção ou inserção de centenas de pares de bases) que podem ser notadas no mRNA.

3.3 Análises de Proteína

Uma vez que muitas aplicações de CRISPR resultam em alterações na *sequência primária e/ou tamanho da cadeia polipeptídica e/ou abundância da proteína*-alvo, técnicas sensíveis a essas modificações são muito úteis (*western blot*, ensaio enzimático, ELISA, etc.).

Contudo, note o leitor que diferentes edições genéticas resultam em diferentes alterações proteicas e, portanto, o pesquisador deve ponderar sobre qual técnica melhor se adequa para detectar a alteração específica promovida na proteína. Por exemplo, ELISA utilizando anticorpos monoclonais capazes de detectar epítomos na porção N-terminal da proteína-alvo podem não ser úteis se o sítio de edição foi na C-terminal e não promoveu a desestabilização da proteína. Devido ao imenso número de possibilidades de edição versus variáveis na análise de proteínas, torna-se inviável discorrer sobre todas elas. Apesar do *western blot* ser a escolha mais comum e útil (por detectar alterações de tamanho e abundância), um estudo detalhado da natureza planejada para a edição genética (éxon-alvo versus a localização do epítopo; KO versus KI; anticorpos monoclonais versus policlonais) permitirá ao pesquisador escolher a melhor opção entre as diversas outras técnicas disponíveis.

4. Cuidados especiais

4.1 Técnicas para identificação de *off-targets*

Uma das principais preocupações em estudos de edição genética são os *off-targets*. Como determinar se outros sítios foram alvejados? Esta questão é de extrema relevância, pois o nocauteamento de outros genes pode comprometer seriamente as conclusões do estudo.

Uma forma de resolver isso é amplificar e sequenciar (após o experimento de CRISPR) os potenciais *off-targets* previamente identificados nas análises *in silico*. Contudo, esta análise é enviesada, levando-nos a avaliar apenas uma pequena parte do genoma, deixando de considerar eventuais *off-targets* não preditos pelos algoritmos disponíveis atualmente. Para realizar análises não enviesadas (*unbiased*) foram desenvolvidas técnicas para aferir a edição de eventuais sítios secundários. Entre elas destacamos: WGS/WES, Digenome-Seq, ChIP-Seq, Guide-Seq, BLESS e HTGTS.

Uma estratégia direta e que usa “força bruta” é o sequenciamento completo do genoma (*whole genome sequencing*, **WGS**) ou do exoma (*whole exome sequencing*, **WES**) com posterior procura *in silico* por sítios apresentando *indels* (Feng et al., 2014; Cho et al., 2014). Nesses dois casos, a edição ocorre *in vivo* e o sequenciamento é realizado a partir das amostras experimentais.

Alternativamente, pode-se extrair o material genético do seu tecido, submetê-lo a clivagem por Cas9/gRNA *in vitro* e posteriormente sequenciar. De acordo com os autores, por meio dessa estratégia denominada **Digenome-Seq** (*sequencing in vitro nuclease-digested genome*) é possível identificar *off-targets* que ocorram em frequências tão baixas quanto 0,1% (Kim et al., 2015). Apesar de o desenvolvimento técnico atual possibilitar tais abordagens, estratégias baseadas exclusivamente em sequenciamento de larga escala não são focadas nos eventos pontuais de edição nem mesmo viáveis para todos os laboratórios.

Por sua vez, Kescu e colaboradores (2014) desenvolveram uma abordagem mais sofisticada baseada em **ChIP-Seq** (*chromatin immunoprecipitation and high-throughput*

Ensaio com dCas9

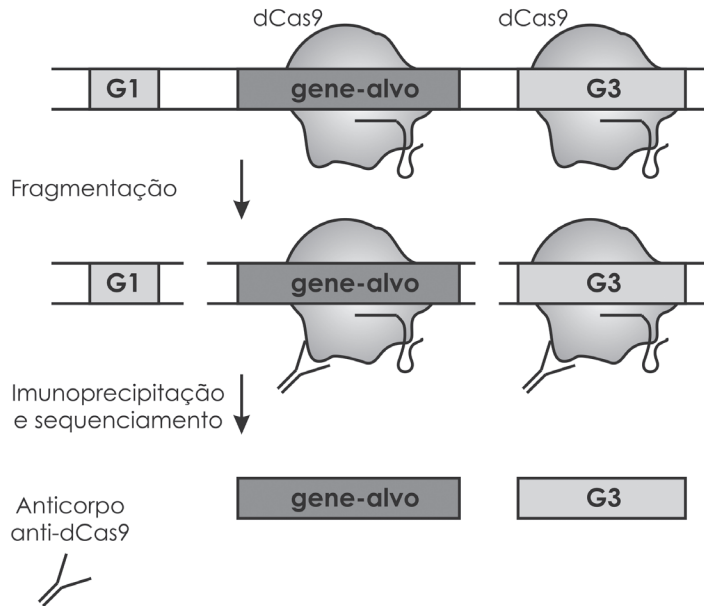


Figura 3. Identificação molecular de *off-targets*: o princípio de ChIP-Seq. Nesta estratégia, o genoma é incubado com dCas9/gRNA e posteriormente submetido a fragmentação. As regiões do genoma interagindo com esse complexo são imunoprecipitadas com o uso de um anticorpo anti-dCas9. Esta amostra de DNA enriquecida é então sequenciada, permitindo a identificação do alvo e dos eventuais *off-targets*. G1 e G3: genes 1 e 3, respectivamente. Imagem: Pereira, TC (2016).

sequencing). Eles incubaram *in vitro* o genoma de interesse com dCas9 associada ao gRNA. A imunoprecipitação destes complexos, seguida de sequenciamento, permite identificar os sítios de interação da nuclease/gRNA com o DNA (figura 3). Contudo, estudos recentes têm apontado que nem todos os *off-targets* evidenciados pela interação com dCas9 são de fato clivados *in vivo* por Cas9 (Tsai et al., 2015; revisto em Lee et al., 2016).

Mali e colaboradores (2013) fusionaram dCas9 a ativadores transcricionais e em seguida realizaram RNA-seq para comparar o perfil transcricional da amostra experimental com um grupo controle. Como esperado, o gene-alvo apresenta um acúmulo de transcritos, ao passo que outros genes que se apresentam induzidos no grupo experimental são interpretados como eventos espúrios de interação de dCas9-ativador/gRNA com *off-targets*.

Tsai e colaboradores (2015) desenvolveram um método denominado **GUIDE-seq** (*genome-wide, unbiased identification of DSBs enabled by sequencing*), que visa identificar sítios de quebra de dupla fita (DSBs). Para isso, administra-se nas células submetidas a edição um pequeno dsDNA de 34 pb, de extremidades abruptas, com ligações do tipo *phosphothiorate* nas extremidades 3' e 5' de ambas as cadeias. Os pesquisadores notaram que essas moléculas se integram em ~33% dos sítios de quebra. Uma vez que se conhece a natureza desses oligonucleotídeos, é possível aplicar estratégias para o sequenciamento dos sítios de integração (Tsai et al., 2015). Assim, regiões do genoma que apresentam DSBs, seja por atividade específica do gRNA, seja por interação espúria (*off-target*) ou até mesmo eventos de outra natureza (radiação), são identificados (figura 4).

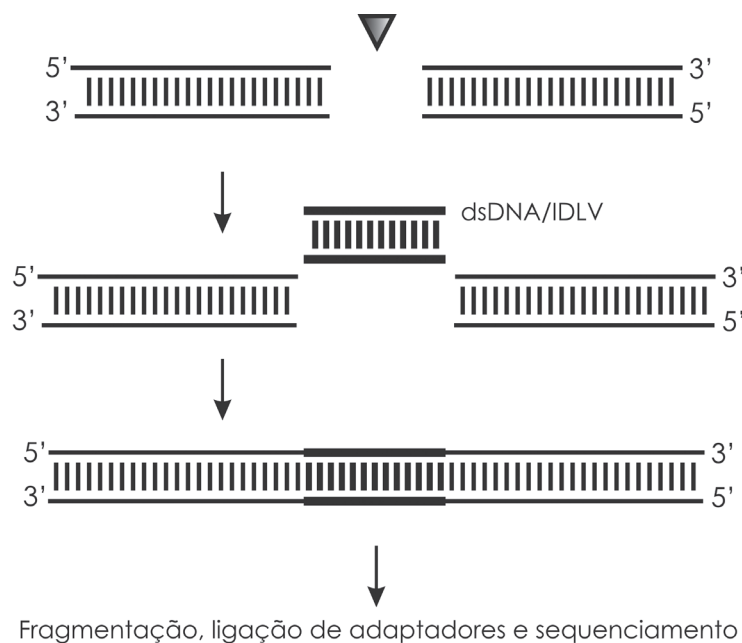


Figura 4. Identificação molecular de *off-targets*: o princípio de Guide-Seq/IDLV. Nessa estratégia, administra-se conjuntamente à Cas9/gRNA uma pequena molécula especial de DNA (ilustrado em preto), que apresenta elevada tendência a se integrar em sítios de DSBs. Uma vez que a sequência nucleotídica desse DNA especial é conhecida, torna-se possível identificar todos os sítios em que ele se integrou, revelando assim eventuais *off-targets*. Triângulo invertido: clivagem mediada por Cas9. Imagem: Pereira, TC (2016).

Uma estratégia conceitualmente semelhante à anterior foi elaborada por Wang e colaboradores (2015). Entretanto, a molécula utilizada para a integração no sítio de DSB é o *vetor lentiviral defeutivo para integrase (IDLV)* (figura 4). De acordo com os autores, esta molécula é preferencialmente integrada em sítios de DSBs via NHEJ, sendo detectada mesmo em sítios *off-targets* cujas frequências sejam tão baixas quanto 1%. Novamente, os sítios de integração podem ser mapeados, uma vez que a sequência nucleotídica do IDLV é conhecida.

Crosseto e colaboradores (2013) desenvolveram uma técnica denominada **BLESS** (*direct in situ breaks labelling, enrichment on streptavidin and next-generation sequencing*), que permite, literalmente, a captura das extremidades de DSBs e posterior sequenciamento. Em síntese, pequenos adaptadores de DNA, fusionados a biotina, são ligados às extremidades livres das DSBs. Em seguida o DNA é fragmentado, e os segmentos de interesse são capturados por avidina e sequenciados, permitindo assim a identificação dos sítios de quebra. Esta estratégia foi aplicada por Ran e colaboradores (2015) para estudar a especificidade de Cas9, identificando eventuais *off-targets* (figura 5).

Por fim, Frock e colaboradores (2015) se basearam no fato de que duas ou mais DSBs simultâneas no genoma (uma no gene-alvo e as eventuais DSBs nos *off-targets*) podem resultar em eventos de translocação. Uma vez que essas translocações envolvem uma sequência conhecida (a do gene-alvo), é possível identificar esses eventos por meio de uma variação de uma técnica denominada **HTGTS** (*high-throughput, genome-wide, translocation sequencing*). Em síntese, a LAM-PCR HTGTS permite amplificar esses sítios de translocação e sequenciá-los.

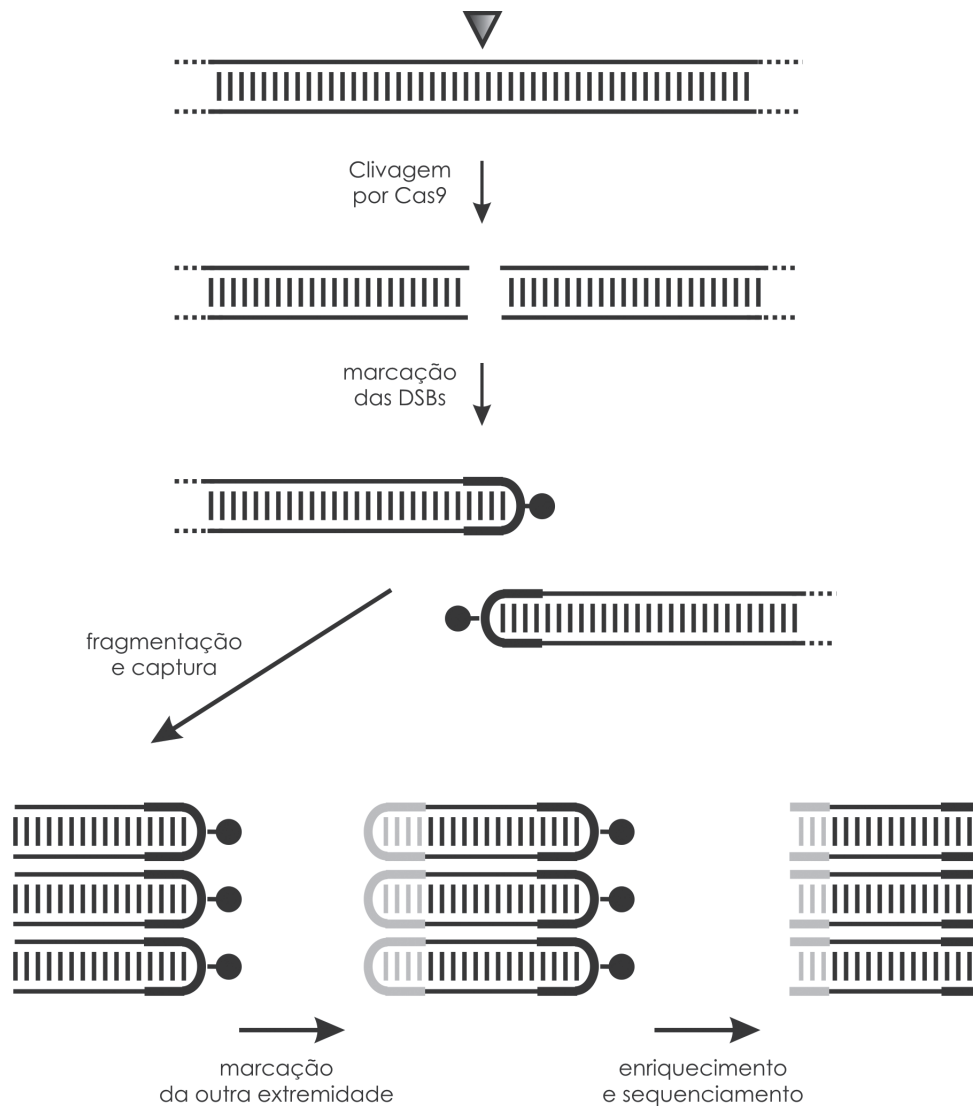


Figura 5. Identificação molecular de *off-targets*: o princípio de BLESS. Nesta estratégia, os sítios de quebra de dupla fita (DSBs) são marcados com um adaptador com biotina (círculo em preto), o DNA é fragmentado e posteriormente capturado por avidina. Após a ligação de um segundo adaptador (cinza-claro), o DNA pode ser sequenciado, revelando os eventuais *off-targets*. Imagem: Pereira, TC (2016).

Tabela 3. Técnicas moleculares para identificação de *off-targets*.

Técnica	Princípio
WGS/WES	Edição seguida por sequenciamento em larga escala
Digenome-Seq	Clivagem <i>in vitro</i> seguida por sequenciamento em larga escala
BLESS	Clivagem seguida por ligação de adaptadores, enriquecimento e sequenciamento
ChIP-Seq	Interação de dCas9/gRNA <i>in vitro</i> , fragmentação, imunoprecipitação e sequenciamento em larga escala
Guide-Seq	Clivagem seguida pela integração de dsDNA, amplificação dos sítios de integração e sequenciamento
HTGTS	Amplificação de eventos de translocação e sequenciamento
dCas9-ativadores transcricionais	Interação de dCas9/gRNA-ativador <i>in vivo</i> , alteração da expressão gênica, RNA-Seq

Lembrando que a translocação resulta em fusões gene-alvo/*off-target*, o sequenciamento desses eventos permite a identificação os sítios espúrios de clivagem. Uma síntese dos princípios de cada técnica pode ser vista na Tabela 3.

4.2 Efeito mosaico

Outra questão frequentemente abordada na edição genômica de precisão em organismos multicelulares, como mamíferos, é o mosaicismo. Mosaicos são organismos que apresentam duas ou mais linhagens celulares somáticas geneticamente distintas, mas todas elas derivadas de um mesmo embrião (figura 6); as quimeras, por sua vez, são derivadas de dois embriões.

Durante o desenvolvimento cada célula sofre mitoses, derivando em duas células-filhas geneticamente idênticas. Se ocorrer uma mutação em uma célula, as divisões subsequentes irão transportar a mutação da sua progenitora. As células mutantes podem, por conseguinte, proliferar em paralelo com as células de tipo selvagem, resultando em um organismo com uma mistura de células com e sem a mutação, ou seja, um organismo mosaico.

De forma geral, quanto mais cedo uma mutação ocorrer durante o desenvolvimento embrionário, maior será o número de células e tecidos mutados no organismo adulto. De maneira oposta, quanto mais tardiamente uma mutação surgir, menor será o número de células/tecidos afetados no adulto. O padrão de distribuição dos tecidos com diferentes composições genéticas em um mosaico pode, portanto, variar muito dependendo da época de ocorrência (figura 6).

Devido ao efeito mosaico, as diversas técnicas que objetivam a criação de animais geneticamente modificados (transgênicos, nocautes, etc.) geralmente não conseguem obtê-los na ‘geração zero’ (F_0 , *i.e.*, o organismo resultante do desenvolvimento do embrião geneticamente manipulado). Tipicamente, esses organismos são mosaicos, contêm um número variável de alelos modificados (ou cópias do transgene) e não necessariamente apresentam o fenótipo esperado. Torna-se, portanto, oportuna a análise de *segregantes* na F_1 e eventualmente cruzamentos controlados até se obter um organismo homozigoto e não mosaico.

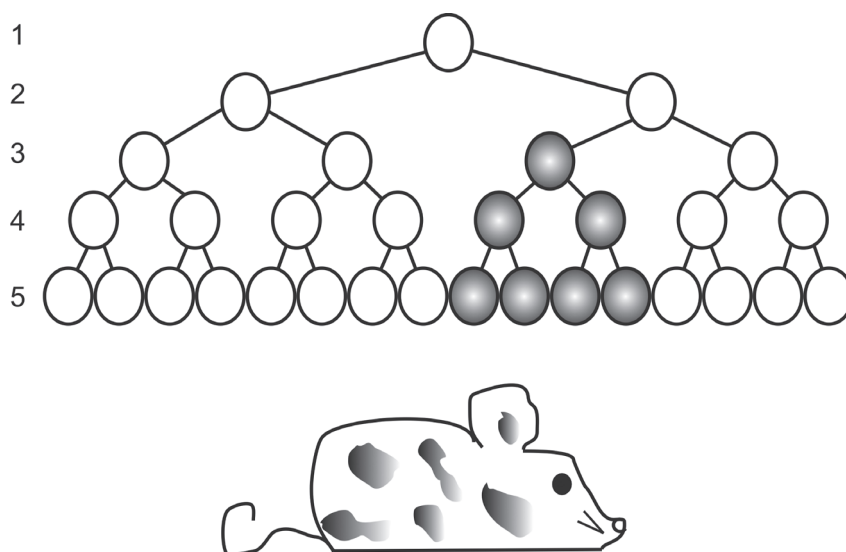


Figura 6. A problemática do mosaicismo na F_0 . Mosaicos são organismos que apresentam duas ou mais linhagens somáticas geneticamente distintas, todas elas derivadas de um mesmo zigoto. A figura ilustra células (círculos) gerando células-filhas por mitose. Quando uma mutação surge em determinada célula (círculo preenchido), toda a linhagem celular dela derivada perpetua essa modificação, gerando tecidos alterados. Essas células/tecidos com a mutação podem ocupar diferentes nichos no organismo adulto (manchas cinza no camundongo). Os números à esquerda simbolizam as gerações celulares. Imagem: Pereira, TC (2016).

Naturalmente, a questão do mosaicismo na F_0 também é observada em experimentos utilizando CRISPR. Contudo, ela pode ser mais ou menos problemática, dependendo da estratégia a ser utilizada. Como exemplo, considere a administração em um zigoto do complexo Cas9/gRNA no formato de uma ribonucleoproteína (RNP) (Kim et al., 2014; Chen et al., 2016). Uma imensa quantidade de situações diferentes pode ocorrer, gerando animais de naturezas distintas (Yen et al., 2014).

Por exemplo, se ainda durante este estado de uma única célula ocorrer a edição de apenas um dos alelos, o animal resultante não será um mosaico, mas sim um heterozigoto. Se porventura a mutação for dominante, o animal F_0 já apresentará o fenótipo.

Por outro lado, se os dois alelos forem editados, novamente ele não será um mosaico, mas sim um *heterozigoto composto*: portando dois alelos mutantes distintos. Isso ocorre porque é muito improvável que a clivagem dos dois alelos resulte na mesma mutação (lembre-se que NHEJ promove *indels* de maneira aleatória). Heterozigotos compostos podem ou não apresentar o fenótipo esperado, dependendo de como é a interação entre estes alelos. Se ambos forem nulos, o animal F_0 já apresentará o fenótipo.

Contudo, imagine o seguinte cenário, que de fato é muito comum: apesar de a administração da RNP ter sido feita no zigoto, a edição pode ter ocorrido apenas no estágio de duas células, sendo apenas uma delas modificada, gerando dois alelos mutantes distintos. Este animal será um mosaico (com células selvagens e células editadas), sendo que as células editadas são heterozigóticas compostas. A análise molecular e fenotípica deste animal F_0 deve considerar todos esses elementos complicadores. É devido a essa problemática que animais F_0

editados por CRISPR são ainda mais complexos do que os mosaicos típicos, aconselhando-se realizar as análises na F_1 ou F_2 .

Entretanto, em alguns casos, organismos F_0 editados por CRISPR podem sim ser homozigotos mutantes, não mosaicos e consequentemente apresentarem o fenótipo esperado. Isto é possível se a abordagem utilizada for a da MCR (Gantz and Bier, 2015) (vide cap. 10).

5. Conclusões

Provavelmente a preocupação central atual quanto ao uso de CRISPR se refere a potenciais efeitos *off-target*. Diversas adaptações têm sido sugeridas para resolver esta questão (revisto em Zhang et al., 2015), como o uso de versões truncadas de gRNAs ou menor concentração final de Cas9/gRNA nas células. Contudo, apesar destas estratégias reduzirem os efeitos *off-target*, elas também diminuem a clivagem *on-target* (dos alvos).

Outros estudos têm sugerido o uso de: (i) Cas9 de alta fidelidade (e.g., SpCas9-HF1 e eSpCas9), (ii) pares de *nickases* ou (iii) de dCas9 fusionadas a domínios FokI (fCas9), todas elas apresentando alta especificidade sem comprometer a edição do alvo (vide cap. 6) (revisto em Zhang et al., 2015).

Uma vez que a criação de linhagens geneticamente modificadas (células ou organismos) é um dos principais objetivos do uso de CRISPR, ter certeza de que outros genes não foram alterados é crucial. Como visto neste capítulo, diversas estratégias *in silico* e *in vitro* para avaliar esta questão, assim como nucleases de elevada precisão para uso *in vivo*, estão disponíveis.

Referências

- Chen S, Lee B, Lee AY, Modzelewski AJ, He L. Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. *J Biol Chem*. 2016;291(28):14457-67.
- Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014;24(1):132-41.
- Crosetto N, et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods*. 2013;10:361-5.
- Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang DL, et al. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc Natl Acad Sci USA*. 2014;111(12):4632-7.
- Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol*. 2015;33(2):179-86.
- Gantz VM, Bier E. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Proc SPIE Int Soc Opt Eng*. 2015;73:389-400.
- Hruscha A, Schmid B. Generation of zebrafish models by CRISPR/Cas9 genome editing. *Methods Mol Biol*. 2015;1254:341-50.
- Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*. 2015;12(3):237-43.
- Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*. 2014;24(6):1012-9.
- Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. 2014;32(7):677-83.
- Lee CM, Cradick TJ, Fine EJ, Bao G. Nuclease Target Site Selection for Maximizing On-target Activity and Minimizing Off-target Effects in Genome Editing. *Mol Ther*. 2016;24(3):475-87.
- Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013;31(9):833-8.
- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015;520(7546):186-91.
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014;32:347-55.

- Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol.* 2015;33:187-97.
- Vouillot L, Thelie A, Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3.* 2015;5(3):407-15.
- Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol.* 2015;33(2):175-8.
- Yang Z, Steentoft C, Hauge C, Hansen L, Thomsen L, Niola F, et al. Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res.* 2015;1-8.
- Yen ST, Zhang M, Deng JM, Usman SJ, Smith CN, Parker-Thornburg J, et al. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Dev Biol.* 2014;393:3-9.
- Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids.* 2015;4:e264.

Métodos para o uso de CRISPR

Capítulo

9

Marcelo Bassalo^{1,2,a}, Gleidson Silva Teixeira^{3,4,b}, Gonçalo Amarante Guimarães Pereira^{3,4,b}, Andréia Prata Vieira^{5,6,c}, Marie Anne Van Sluys^{6,c}, Joan Barau^{7,d}, Kaio Cesar Simiano Tavares^{8,d}, Juliana Borsoi^{9,e} e Lygia da Veiga Pereira^{9,e}

¹Dept. of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, CO - EUA

²Dept. of Chemical and Biological Engineering, University of Colorado Boulder, CO - EUA

³Dept.º de Genética Evolução e Bioagentes, IB, UNICAMP, SP - Brasil

⁴GranBio Celere, SP - Brasil

⁵Dept.º de Botânica da Universidade de São Paulo - IB, USP, SP - Brasil

⁶Dept. of Plant Biology and Pathology, SEBS, The State University of New Jersey, NJ - EUA

⁷Institut Curie, Paris Sciences et Lettres Research University, Paris - França

⁸Lab. de Biologia Molecular e do Desenvolvimento, NBE, UNIFOR, CE - Brasil

⁹Dept.º de Genética e Biologia Evolutiva, IB, USP, SP - Brasil

^aAutoria da seção: 2

^bAutoria das seções: 3 e 4

^cAutoria da seção: 5

^dAutoria da seção: 6

^eAutoria da seção: 7

Estruturação do capítulo

1. Introdução
2. Bactérias
 - 2.1 Introdução
 - 2.2 Regras para desenho e construção de gRNAs e DNA doador para bactérias
 - 2.3 Edição da sequência nucleotídica
 - 2.4 Nocauteamento e deleções gênicas
 - 2.5 Inserção de genes heterólogos
3. Leveduras
 - 3.1 Introdução
 - 3.2 Nocauteamento de genes
 - 3.3 Integração de transgenes
 - 3.4 Integração de vias metabólicas
4. Fungos filamentosos
 - 4.1 Introdução
 - 4.2 Exemplos
5. Plantas
 - 5.1 Introdução
 - 5.2 Inserção de transgene (KI) em *Arabidopsis*
 - 5.3 Deleção de sequências genômicas em arroz
 - 5.4 Geração de variabilidade genética livre de DNAs exógenos
 - 5.5 Regulação da expressão gênica *in planta*
 - 5.6 Próximos passos do uso de CRISPR em plantas
6. Animais
 - 6.1 *Drosophila melanogaster* (Chen et al., 2016)
 - 6.2 *Caenorhabditis elegans* (Paix et al., 2014)
 - 6.3 *Xenopus laevis* (Wang et al., 2015)
 - 6.4 *Zebrafish* (Di Donato et al., 2016)
 - 6.5 Camundongos (Maddalo et al., 2014)
 - 6.6 Ratos (Bakondi et al., 2016)
 - 6.7 Cabras (Ni et al., 2014)
 - 6.8 Bovinos (Heo et al., 2015)
 - 6.9 Cães (Zou et al., 2015)
 - 6.10 Suínos (Yang et al., 2015)
 - 6.11 Macacos (Chen et al., 2015; Kang et al., 2015)
 - 6.12 Conclusões
7. Células humanas
 - 7.1 Introdução
 - 7.2 Substituição alélica
 - 7.3 Deleção de sequências gênicas
 - 7.4 Inativação gênica
 - 7.5 Integração de sequências exógenas (*knock-in*)
 - 7.6 Edição genética em espermatozônios

1. Introdução

Neste capítulo apresentaremos detalhes técnicos da aplicação de CRISPR em uma imensa variedade de modelos, com diversas aplicações distintas, para que o leitor possa ter uma visão de possíveis estratégias para utilizar em sua espécie de pesquisa.

2. Bactérias

2.1 Introdução

Desde as primeiras demonstrações da utilização de CRISPR para edição de genomas (Cong et al., 2013; Jiang et al., 2013; Jinek et al., 2012; Mali et al., 2013), o uso da técnica se expandiu para diversos grupos de organismos e hoje inclui bactérias, fungos, plantas, insetos, mamíferos, entre outros (Sander and Joung, 2014). No entanto, a aplicação da técnica em bactérias é relativamente mal representada em comparação com eucariotos, existindo apenas poucos estudos que empregaram a metodologia neste grupo até então. Dessa forma, a edição de genomas de bactérias utilizando CRISPR é uma área ainda a ser muito explorada e com grandes potenciais aplicações para a biotecnologia e microbiologia.

Em comparação com eucariotos, a utilização de CRISPR em bactérias se difere devido aos mecanismos de reparo de DNA. Ao se introduzir uma quebra de fita dupla no genoma, eucariotos se utilizam de duas vias principais para reparar a lesão: HDR e NHEJ (vide cap. 2). No entanto, a maioria das espécies de bactérias não possui uma via NHEJ, ou possui uma via ineficiente, *resultando em uma letalidade causada pela CRISPR após a quebra de fita dupla* (Gomaa et al., 2014; Jiang et al., 2013). Dessa forma, a utilização de CRISPR em bactérias é baseada majoritariamente na via HDR, com um molde homólogo sendo fornecido para reparar a quebra de dupla fita. Tendo em vista a ausência da via NHEJ, um importante aspecto da utilização da técnica em bactérias é o fato de que CRISPR atua como um mecanismo de seleção por genomas editados, uma vez que células que não incorporarem o DNA doador (*i.e.*, o **molde homólogo**) pela via HDR sofrem morte celular devido à lesão no genoma (figura 1a).

Nesta seção, serão introduzidas algumas regras para desenho de gRNAs e moldes para HDR, visando a edição de genomas bacterianos. Em seguida, protocolos que utilizaram CRISPR para diferentes fins em bactérias serão descritos para exemplificar a ampla aplicabilidade da técnica.

2.2 Regras para desenho e construção de gRNAs e DNA doador para bactérias

Apesar de alguns estudos utilizarem outros sistemas de CRISPR em bactérias (Vercoe et al., 2013), as regras e protocolos aqui descritos focam nas aplicações com a endonuclease Cas9 de *Streptococcus pyogenes* (sistema Tipo II). Tal sistema requer a presença da sequência PAM (NGG) a 3' da sequência-alvo. Com isso, o primeiro passo é escolher a sequência-alvo de 20 nucleotídeos no genoma da bactéria, a qual deve conter imediatamente à 3' a sequência PAM (figura 1b). Diversas estratégias computacionais estão disponíveis para auxiliar na escolha de sequências com maiores chances de atividades específicas (Doench et al., 2014; Heigwer et al., 2014). No entanto, é importante destacar que clivagens não específicas (*off-targets*)

a

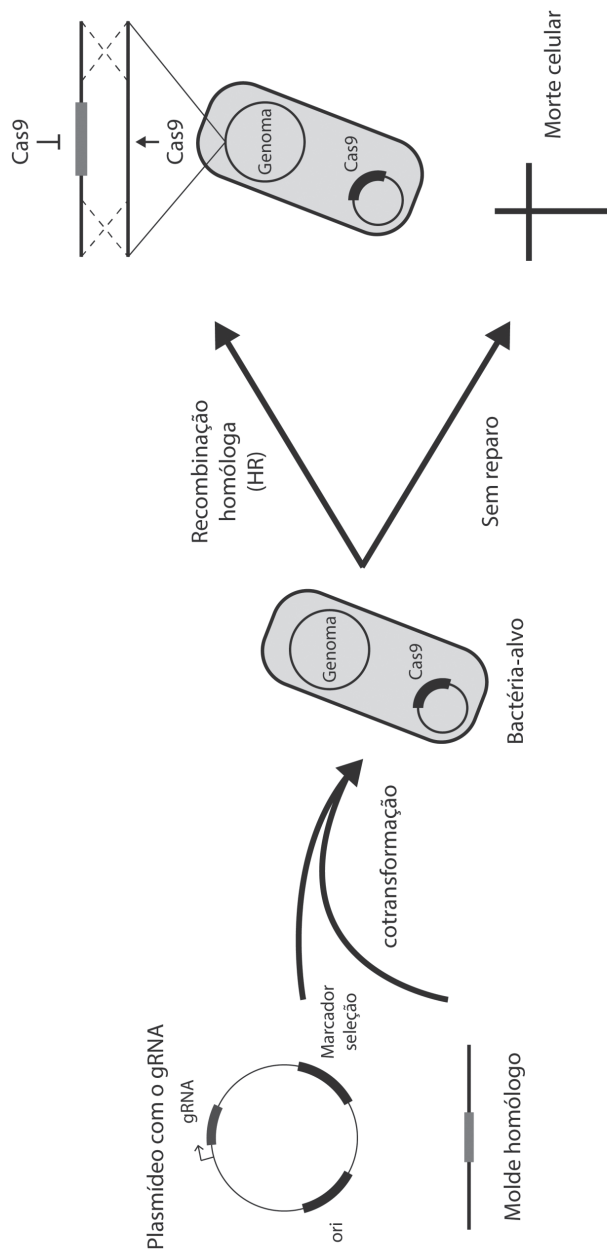


Figura 1a. Lógica da utilização de CRISPR em bactérias. Os protocolos geralmente se baseiam na cotransformação de um gRNA que tem como alvo o genoma juntamente com um molde homólogo para reparo. Devido à ausência ou ineficiência da via NHEJ, células que não incorporarem o molde homólogo (DNA doador) pela via HDR sofrem morte celular; selecionando-se por genomas editados. Imagem: Bassalo, M; Teixeira, GS e Pereira, GAG (2016).

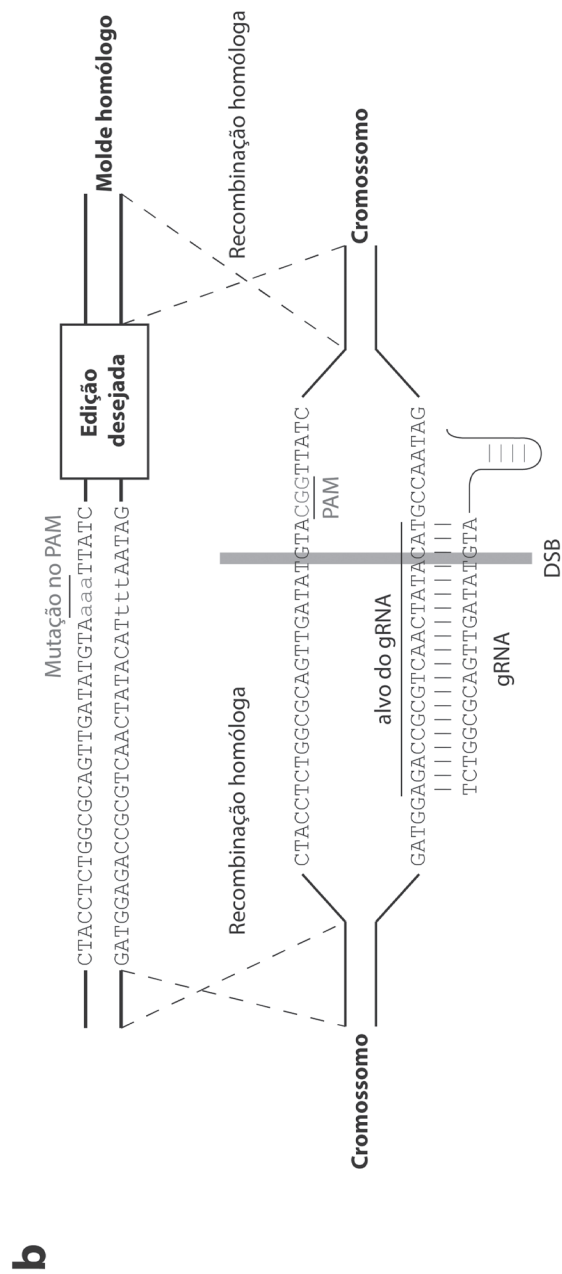


Figura 1b. Lógica da utilização de CRISPR em bactérias. Regras para o desenho do molde homólogo e do gRNA para edição de genomas de bactérias. DSB: quebra de fita dupla. Imagem: Bassalo, M; Teixeira, GS e Pereira, GAG (2016).

são eliminadas da população em bactérias, uma vez que o reparo pela via NHEJ é ineficiente ou inexistente.

Uma vez escolhidas as *sequências-alvos* de 20 nucleotídeos (também conhecidas como *spacers*), estas podem ser facilmente incluídas em oligonucleotídeos para clonagem em vetores contendo o *scaffold* do gRNA sob controle de um promotor da bactéria a ser editada (diversas opções disponíveis em banco de dados como o Addgene). Com isso, a sequência *spacer* desses vetores pode ser substituída pela sequência *spacer*-alvo contida nos oligonucleotídeos através de metodologias de clonagem baseadas em homologia como Gibson (Gibson et al., 2009) ou CPEC (Quan and Tian, 2011), as quais podem ser facilmente realizadas em *multiplex* para construção de bibliotecas. É importante destacar que a clonagem dos gRNAs não deve ser feita em um vetor contendo a Cas9 na bactéria que possui o alvo a ser editado. Isso porque a expressão basal da Cas9 e dos gRNAs é suficiente para clivar o genoma da bactéria, resultando em: i) morte celular caso o DNA doador não esteja presente; ii) enriquecimento de células que clonaram um vetor não funcional (sem clivagem no genoma, menor toxicidade). Dessa forma, a clonagem dos gRNAs deve ser feita na ausência da Cas9 (vetores separados) ou em espécies de bactérias que não possuam o alvo da edição, de forma que a combinação da Cas9 e gRNA só ocorra na célula-alvo no momento desejado para edição do genoma.

O DNA doador para edição deve ser construído de forma a conter a *nova sequência* (sequência endógena com alguma mutação ou um transgene) flanqueada por regiões de homologia à região-alvo no genoma (figura 1b). O tamanho da região de homologia varia de acordo com a espécie de bactéria (ou seja, com a eficiência do sistema HDR) e com o tamanho da *nova sequência* que está sendo introduzida. É importante destacar que o DNA doador deve ser construído de forma a *inutilizar* a sequência PAM no genoma, tornando as células editadas imunes à toxicidade de CRISPR (*i.e.*, de forma a impedir sucessivas clivagens no genoma mesmo após a incorporação do DNA doador). Isso pode ser obtido através da deleção da sequência PAM ou através da mutação desta nos braços de homologia (figura 1b). Por fim, os moldes homólogos construídos podem ser fornecidos às células para edição tanto em forma linear (produto de PCR ou digestão) quanto circular (plasmídeos), dependendo da bactéria a ser editada.

2.3 Edição da sequência nucleotídica

Diversas aplicações biotecnológicas requerem o desenvolvimento de linhagens de bactérias que apresentam elevada produtividade e resistência a condições industriais. Isso geralmente é obtido através de modificações como o redirecionamento do fluxo de nutrientes, a otimização da expressão de genes ou da atividade de proteínas específicas. Com isso, a capacidade de editar elementos específicos no genoma de bactérias (promotores, sítios de ligação de ribossomos - RBS, regiões codificadoras, etc.) vem sendo buscada há anos pela área de engenharia metabólica. A maioria dos métodos disponíveis para realizar tais modificações específicas se baseia na utilização de oligonucleotídeos que são incorporados no genoma através de recombinases endógenas ou exógenas. Este processo é conhecido como “*recombineering*” e possui baixa eficiência (Sharan et al., 2009; Wang et al., 2009). No entanto, a combinação do processo de *recombineering* junto com a seleção feita por CRISPR abre portas não somente para aumentar a eficiência substancialmente, mas também para viabilizar o processo em bactérias de difícil manipulação genética.

Como exemplo, bactérias lácticas têm grande importância na indústria química e farmacêutica, sendo a edição de seus genomas, portanto, essencial. Como prova de conceito de

que CRISPR pode ser aplicado em espécies que não são modelos, a bactéria láctica *Lactobacillus reuteri* foi editada através da combinação de CRISPR com *recombineering* (Oh and van Pijkeren, 2014). Para tal, os autores se basearam em um sistema de CRISPR em que o crRNA e o tracrRNA são transcritos separadamente, em vez de fusionados como um gRNA. Os autores transformaram *L. reuteri* com dois plasmídeos: um contendo a Cas9 e o tracrRNA, e outro contendo recT, uma recombinase que viabiliza o processo de *recombineering*. Edições no genoma podem então ser executadas através da cotransformação com um oligonucleotídeo para reparo junto com um terceiro plasmídeo contendo o crRNA que tem como alvo o genoma.

Para confirmar a introdução das mutações no genoma, foi utilizada uma metodologia conhecida como MAMA-PCR (*mismatch amplification mutation assay-PCR*). Para isso, um conjunto de *primers* são desenhados de forma que alguns deles permitem amplificar o fragmento do genoma caso a mutação não esteja presente. Outros *primers* se anelam na mutação inserida, de forma que a amplificação só ocorre caso a mutação seja corretamente inserida no genoma (Cha et al., 1992). Utilizando MAMA-PCR, os autores detectaram eficiências próximas a 100% para mutações pontuais inseridas utilizando CRISPR em *L. reuteri*.

A estratégia usada para editar o genoma de *L. reuteri* pode ser aplicada utilizando oligonucleotídeos degenerados para construção de bibliotecas de mutantes. Para isso, oligonucleotídeos são sintetizados de forma a conter uma mutação silenciosa na sequência PAM e bases degeneradas (NNN) para alteração do códon-alvo. Ao desenhar um conjunto de oligonucleotídeos com o objetivo de introduzir bases degeneradas ao longo dos códons de uma proteína-alvo, é possível construir uma biblioteca de saturação de códons (*site-saturation mutagenesis*) diretamente no genoma da bactéria. Tal capacidade é uma importante ferramenta para engenharia de proteínas, uma vez que bibliotecas de mutantes para proteínas endógenas podem ser facilmente geradas e testadas visando a otimização de sua atividade. Os autores utilizaram essa estratégia com oligonucleotídeos contendo a sequência degenerada NNK visando a construção de uma biblioteca de saturação de códons. A sequência NNK permite a mutação para 32 códons codificando todos os 20 aminoácidos. Ao sequenciar 91 colônias, os autores identificaram 5 que escaparam de CRISPR no processo e 86 que apresentaram mutações para 19 outros aminoácidos, sugerindo uma alta eficiência para o método.

A edição de genomas bacterianos utilizando CRISPR foi também recentemente demonstrada de forma combinatória utilizando a bactéria *Escherichia coli* (Ronda et al., 2016). O protocolo permite a realização de sucessivos ciclos (*rounds*) de edição visando a construção de combinações de mutações no genoma, o que é obtido em um sistema compreendido por dois plasmídeos. O primeiro contém a Cas9 e a maquinaria necessária para realização do *recombineering* através do sistema red do fago lambda. O segundo é um vetor a ser reutilizado a cada ciclo de edição. Este último possui o gRNA que tem como alvo o genoma, além de um *array* de CRISPR (crRNA separado do tracrRNA) que tem como alvo o próprio plasmídeo.

Dessa forma, as edições podem ser feitas através da cotransformação com moldes homólogos na forma de oligonucleotídeos junto com o gRNA clonado no segundo vetor. Após a edição do genoma, um tratamento com ramnose e anidrotetraciclina induz o *array* de CRISPR e resulta na eliminação (*cura*) do plasmídeo contendo o gRNA, permitindo um novo ciclo de transformação e edição na linhagem. Utilizando esse protocolo, os autores demonstraram a edição de três sítios-alvos no genoma (genes *galK*, *lacZ* e *xylA*, envolvidos no metabolismo de açúcares). Para identificar a eficiência de edição, os autores introduziram códons de terminação nos três genes-alvos. Com isso, colônias que incorporaram a mutação podem ser facilmente identificadas através de um sistema de *screening* baseado em meio

MacConkey ágar, no qual colônias selvagens consomem o açúcar do meio e assumem uma coloração avermelhada. As colônias mutantes que incorporaram o códon de terminação são incapazes de consumir o açúcar correspondente suplementado no meio e assumem uma coloração branca. Utilizando esse ensaio, o protocolo permitiu a edição dos três sítios-alvos (*galK*, *lacZ* e *xylA*) com eficiência entre 96 e 99,7%. Devido à alta eficiência do protocolo, os autores também conseguiram editar dois genes simultaneamente, abrindo portas para geração de bibliotecas combinatórias em *multiplex*.

2.4 Nocauteamento e deleções gênicas

O nocauteamento (KO) de genes em bactérias é outra ferramenta importante tanto para a biotecnologia quanto para a microbiologia. Deleções e silenciamentos são frequentemente utilizados para redirecionar o fluxo de nutrientes e maximizar a produção em linhagens industriais. Além disso, a geração de KOs é uma clássica ferramenta genética para descobrir a função de novos genes. Dessa forma, uma das primeiras ferramentas genéticas desejadas para a manipulação de espécies não modelos de bactérias é a capacidade de gerar KOs.

Protocolos utilizando CRISPR foram desenvolvidos para geração de KOs com alta eficiência em espécies que não dispõem de amplas ferramentas para manipulação genética, como espécies do gênero *Streptomyces* (Cobb et al., 2015). Esse grupo tem ampla importância devido à presença de inúmeros *clusters* de genes (muitos ainda com função desconhecida) envolvidos com a produção de metabólitos secundários de potencial aplicação farmacológica. O protocolo é baseado em um único plasmídeo, o qual contém a Cas9 (com códons otimizados para *Streptomyces*, devido ao alto conteúdo GC encontrado no gênero), um sítio para clonagem do gRNA, um sítio para clonagem do molde homólogo e uma origem de replicação sensível a temperatura (o que permite eliminar o plasmídeo após a edição do genoma). Além disso, o plasmídeo possui uma origem de transferência, permitindo realizar a clonagem em *E. coli* e transferir o plasmídeo final por conjugação para a espécie de *Streptomyces* desejada.

Com isso, após a clonagem do gRNA e do DNA doador em *E. coli*, o plasmídeo pode ser transferido para edição de *Streptomyces*. Para a geração de KOs, é utilizado um gRNA que tem como alvo a região a ser deletada, além de um DNA doador contendo 1 kb de homologia à 5' e 3', desenhado de forma que a incorporação dele por HDR resulta na deleção de alguns pares de base no genoma (mutação de mudança no quadro de leitura do gene-alvo). A geração dos KOs alvos foi observada com eficiências próximas a 100% para diferentes alvos em diferentes espécies de *Streptomyces*. Os autores do protocolo também demonstraram que é possível clonar no vetor um conjunto de dois gRNAs com os respectivos DNA doadores, espaçados de forma a deletar regiões maiores do genoma. Utilizando essa dupla edição, foi possível deletar *clusters* de genes completos com tamanho de 31 kb, o que foi confirmado por PCR e a nível fenotípico (ausência do pigmento produzido pelo *cluster* de gene deletado). Dessa forma, o protocolo pode ser utilizado tanto para a geração de KOs quanto para deleções gênicas.

Outro protocolo para o gênero *Streptomyces* desenvolvido permite a geração de KOs e deleções gênicas, de forma similar à descrita acima (Tong et al., 2015). No entanto, os autores desse último protocolo exploraram não somente a geração de KOs pela via HDR, mas também pela via NHEJ. Isso é possível devido ao fato de que bactérias do gênero *Streptomyces* possuem uma via NHEJ incompleta, funcionando de forma ineficiente. Com isso, os autores demonstraram que, com a utilização de apenas um gRNA que tem como alvo o gene a ser editado (sem o DNA doador), a via NHEJ resulta em deleções nucleotídicas de tamanhos

variados, resultando na inativação do gene-alvo. As eficiências observadas variaram entre 3% e 54%, considerando a ineficiência da via NHEJ. Curiosamente, os autores desenvolveram um vetor que reconstitui a via NHEJ incompleta de *Streptomyces* (possível através da expressão da ligase LigD). Dessa forma, com a utilização desse último vetor, é possível gerar KOs com maior eficiência usando apenas um gRNA, sem DNAs doadores.

2.5 Inserção de genes heterólogos

Por fim, a inserção de genes heterólogos (*knock-ins* ou KIs) é outra importante ferramenta na engenharia genética. A inserção de construções no genoma permite não somente a manipulação de funções endógenas, mas também implementar novas funções de forma estável e com níveis de expressão fisiologicamente relevantes. CRISPR também foi utilizado para facilitar a inserção de genes heterólogos no genoma de bactérias. Como exemplo, a inserção de genes e até mesmo vias bioquímicas completas podem ser obtidas com alta eficiência na bactéria *Escherichia coli* (Bassalo et al., 2016).

Para tal, a bactéria é inicialmente transformada com dois vetores: um contendo a Cas9 (induzida por arabinose) e outro que expressa o sistema red do fago lambda para aumentar a eficiência de HDR em *E. coli*, vetor pSIM5 induzido por temperatura (Datta et al., 2006). Com isso, a bactéria contendo os dois vetores é crescida até a fase logarítmica, seguida de uma indução com arabinose e alta temperatura. Após, a inserção de genes heterólogos é possível em uma única etapa através da cotransformação com um terceiro vetor contendo o gRNA junto com o DNA doador para reparo. Este molde homólogo pode ser fornecido tanto de forma linear (produto de PCR) ou clonado em um vetor, embora a forma linear seja mais favorável, pois ele é naturalmente eliminado após a edição do genoma.

Utilizando esse protocolo, a inserção da proteína fluorescente GFP foi possível com eficiências próximas a 100% em diferentes locais do genoma, o que foi observado por microscopia de fluorescência e citometria de fluxo. Os autores demonstraram que a eficiência de inserção se mantém alta utilizando DNAs doadores com braços de homologia desde 50 bp até 600 bp. Devido ao pequeno tamanho, estas sequências homólogas podem ser introduzidas diretamente no gene heterólogo-alvo por PCR, facilitando a construção do DNA doador a ser inserido no genoma. Adicionalmente, os autores demonstraram que vias bioquímicas completas também podem ser inseridas utilizando esse protocolo. Como exemplo, uma via de 10 kb contendo os genes necessários para a produção do biocombustível isobutanol foi inserida em um único dia com 50% de eficiência usando braços de homologia de 100 bp. Dessa forma, o protocolo abre portas para construir e testar novas vias bioquímicas com facilidade diretamente no genoma da bactéria, gerando linhagens estáveis susceptíveis a processos industriais ou a subseqüentes manipulações genéticas.

Portanto, a aplicação de CRISPR em bactérias facilitou significativamente a manipulação genética tanto de espécies-modelos como não modelos. Diversos protocolos estão disponíveis para a edição específica da sequência nucleotídica, geração de deleções e a inserção de construções heterólogas. A alta eficiência observada para as diversas formas de edição em diferentes espécies de bactérias fortalece a ampla aplicabilidade e facilidade da técnica. Com a explosão da utilização de CRISPR em diferentes organismos nos últimos anos, a exploração da técnica em bactérias deve trazer avanços significativos para a microbiologia e biotecnologia nos próximos anos.

3. Leveduras

3.1 Introdução

Por muitas décadas, leveduras têm sido os micro-organismos mais utilizados como modelo experimental para o estudo de aspectos fundamentais da biologia de eucariotos, especialmente as espécies *Saccharomyces cerevisiae* e *Schizosaccharomyces pombe* (Revisado por Botstein & Fink, 2011 e Duina et al., 2014). Com o advento da biotecnologia moderna, entretanto, *S. cerevisiae* e outras espécies como *Picchia pastoris* (*Komagataella pastoris*) e *Kluyveromyces spp.* se tornaram plataformas para o desenvolvimento de “biofábricas” aplicadas à produção industrial de fármacos, bioquímicos, entre outros.

Neste contexto, abordagens de edição do DNA baseadas em CRISPR permitem o estudo rápido da função de genes, inserção ou substituição de mutações pontuais, integração estável de um único gene ou de vias metabólicas completas, além da análise de sinergia entre alelos e sua contribuição para um determinado fenótipo. Em espécies de leveduras consideradas “não modelos”, CRISPR destaca-se como uma alternativa à escassez de ferramentas de manipulação genética adequadas ou à baixa eficiência de recombinação homóloga sítio-dirigida. A funcionalidade da técnica CRISPR já foi demonstrada em *S. pombe* (Jacobs et al., 2014), *K. lactis* (Horwitz et al., 2015) e principalmente em *S. cerevisiae*, sendo esta a espécie preferida para muitas aplicações biotecnológicas, incluindo biocombustíveis, produtos químicos farmacêuticos, etc.

Desde a primeira demonstração da funcionalidade do sistema CRISPR em *S. cerevisiae* (Dicarlo et al., 2013) foram propostas diferentes estratégias com variado impacto sobre a eficiência de edição. Para a expressão da Cas9, prevalece o uso de promotores fortes ou medianos em plasmídeos centroméricos (CEN/ARS) – origem de replicação para poucas cópias por célula. O uso de plasmídeos com origem de replicação 2 μ – múltiplas cópias – combinado a promotores fortes como o TEF1 tem sido associado à toxicidade (Ryan et al., 2014), embora este resultado não tenha sido observado por Bao et al. (2015). Assim como Dicarlo et al. (2013), a maioria das abordagens utilizam a Cas9 derivada de *Streptococcus pyogenes* com códons previamente otimizados para humanos (Bao et al., 2015; Dicarlo et al., 2013; Jakočiūnas et al., 2015a; Ronda et al., 2015; Stovicek et al., 2015; Zhang et al., 2014). Entretanto, foi demonstrada também a funcionalidade das versões original ou otimizada para *S. cerevisiae* (Horwitz et al., 2015; Mans et al., 2015; Ryan et al., 2014).

Em relação ao crRNA e ao componente estrutural tracrRNA, a maioria das abordagens mantém a estratégia de transcrição de um RNA quimérico ou gRNA (crRNA + tracrRNA) dirigida por elementos regulatórios para a RNA polimerase III. Para isto, utiliza-se o promotor SNR52 (snoRNA) e a região 3' do gene SUP4 (tRNA) como terminador (Dicarlo et al., 2013; Horwitz et al., 2015; Jakočiūnas et al., 2015a; Jakočiūnas et al., 2015b; Ronda et al., 2015; Stovicek et al., 2015). Esta combinação produz transcritos com menos de seis resíduos de uracila na extremidade 3' do RNA, consideração importante, uma vez que originalmente o componente estrutural do gRNA contém quatro resíduos (Acker et al., 2013; Mali et al., 2013). Pelo menos duas outras estratégias foram demonstradas e devem ser consideradas caso se pretenda substituir o promotor SNR52 (Ryan et al., 2014) ou transcrever múltiplos gRNAs dispostos de modo semelhante ao loco CRISPR em procariotos (Bao et al., 2015).

A decisão pela melhor estratégia para a aplicação da técnica de CRISPR em *S. cerevisiae* deve levar em consideração o número de edições necessárias, o *background* genético e a

ploidia da linhagem. Neste contexto, são apresentados exemplos de abordagens utilizadas com alta eficiência na construção de *knockouts*, inserção/substituição de mutações pontuais, inserção de genes heterólogos e de vias metabólicas.

3.2 Nocauteamento de genes

Em *S. cerevisiae*, a construção de linhagens *knockout* (KO) ou *knock-in* (KI) baseia-se em sua habilidade de realizar recombinação homóloga. Linhagens são tipicamente transformadas com um DNA doador contendo a sequência de interesse flanqueada por sequências homólogas à região na qual se pretende promover a edição (Storici et al., 2003). Dicarlo et al. (2013) foram os primeiros a utilizar CRISPR para produzir KOs por meio da recombinação homóloga. Como prova de conceito realizaram a edição do loco CAN1, gene que codifica uma permease de membrana cuja perda de função confere resistência à canavanina, um análogo tóxico da arginina (Whelan et al., 1979).

Em resumo, células expressando constitutivamente a Cas9 foram simultaneamente co-transformadas com plasmídeo contendo o gRNA para edição do loco CAN1 e duas versões de DNA doador sendo: (i) oligonucleotídeo de dupla fita com 90 pb de extensão (dsOligo) contendo um códon de terminação (TGA) prematuro; (ii) cassete de expressão KanMX – que confere resistência ao antibiótico geneticina (G418) –, contendo 1,4 mil pb de extensão e 50 pb de homologia ao loco CAN1. Ambos flanqueiam a sequência PAM e geram KOs tolerantes à canavanina.

É importante ressaltar que neste exemplo a Cas9 e o gRNA foram transcritos *in vivo* a partir de plasmídeos distintos (P01-02, Tabela 1), de modo que transformantes foram obtidos pela dupla seleção em meio sintético sem triptofano (seleção para a Cas9) e sem uracila (seleção para o gRNA) (Mumberg et al., 1995). Para quantificar a eficiência das edições, colônias contendo ambos os plasmídeos foram transferidas para meio de cultivo contendo canavanina ou G148. Para ambos os DNAs doadores, aproximadamente 100% das células selecionadas apresentaram resistência à canavanina. A mesma eficiência foi observada quando células cotransformadas com o cassete KanMX foram selecionadas em meio contendo G148, confirmando a inserção do cassete (Dicarlo et al., 2013).

Embora seja altamente eficiente, esta abordagem se limita às linhagens que possuem marcas auxotróficas. No entanto, este sistema pode ser facilmente adaptado para conter em um único vetor ambos os componentes do sistema CRISPR e um marcador de seleção dominante,

Tabela 1. Vetores e elementos regulatórios utilizando em diferentes aplicações de CRISPR.

ID	Elementos regulatórios	Replicação e seleção	Referência
P01	pTEF1-Cas9-tCYC1	CEN/ARS, TRP1	Dicarlo et al. 2013
P02	pSNR52-gRNA-tSUP4	2μ, URA3	Dicarlo et al. 2013
P03	pTEF1-Cas9-tCYC1-pSNR52-gRNA-tSUP4	2μ, KanMX	DNA2.0
P04	pSNR52-gRNA-tSUP4	2μ, LEU2	Jakočiūnas et al. 2015a
P05	pRNR2-Cas9-tCYC1-ptRNA ^{Pro} -HDV-gRNA-tSNR52	2μ, KanMX	Ryan et al. 2014
P06	pTEF1-iCas9-tADH2-pSNR52-DR-gRNA-DR-tSUP4-pRPR1-tracrRNA-tRPR1	2μ, URA3	Shi et al. 2016

como o cassete KanMX (P03, Tabela 1). Em testes preliminares, células cotransformadas com 200 ng deste plasmídeo (2 μ) e do dsOligo como doador foram selecionadas em meio de cultivo padrão contendo 200 μ g/mL de G418. Pelo menos 50% de eficiência de edição foi observada entre os transformantes selecionados, tanto para células haploides quanto para diploides, ambas derivadas da linhagem industrial JAY270 (comunicação pessoal).

Com o objetivo de acelerar estudos de genômica funcional, diversos autores elaboraram estratégias baseadas em CRISPR para edição simultânea de múltiplos locos (Bao et al., 2015; Horwitz et al., 2015; Jakočiūnas et al., 2015a; Ryan et al., 2014; Zhang et al., 2014). Entre elas, Jakočiūnas et al. (2015a) descreveram uma abordagem altamente específica e eficiente para edição de até cinco locos em um único evento de transformação. Neste estudo, os autores utilizaram CRISPR para obter mutantes KO para quatro locos (*BTS1*, *ROX1*, *YPL062W* e *YJL064W*) e *knockdown* para o gene essencial *ERG9*, todos previamente descritos como relacionados à produção de mevalonato, um intermediário-chave para a produção industrial de isoprenoides (Asadollahi et al., 2010; Özaydin et al., 2013).

Em linhas gerais, os autores utilizaram uma abordagem semelhantes à demonstrada por Dicarlo et al. (2013): (i) expressão da Cas9 controlada pelo promotor TEF1 (P01, Tabela 1); dsOligo com 90 pb de extensão como DNA doador, sendo a região PAM substituída por um códon de terminação (exceto para *ERG9*); (iii) pSRN52 e tSUP4 como elementos regulatórios para expressão dos gRNAs (P04, Tabela 1) e; (iv) dupla seleção de transformantes para recuperar células contendo ambos os plasmídeos que carregam a Cas9 e os gRNAs.

A principal modificação inclui a construção de plasmídeos contendo múltiplos gRNAs utilizando o método *USER cloning* descrito por Jensen et al. (2014). Em resumo, cassetes de expressão de cada gRNA foram amplificados utilizando oligonucleotídeos universais compatíveis ao método de clonagem. Em seguida, todas as combinações possíveis de dois, três, quatro e cinco gRNAs foram inseridas no vetor digerido com as enzimas *AsiSI* e *Nb.BsmI* (Jakočiūnas et al., 2015a).

Células expressando constitutivamente a Cas9 foram simultaneamente cotransformadas com cada uma das 31 versões de plasmídeo e DNAs doadores apropriados. Transformantes foram selecionados em meio sintético sem triptofano (seleção para a Cas9) e sem leucina (seleção para o gRNA) e a eficiência de edição observada foi de 50% a 100% para todas as combinações. Em seguida, os autores avaliaram a produção de mevalonato dos 31 mutantes em comparação com a linhagem parental selvagem CEN.PK113-7D. Como resultado, 20 das 31 linhagens apresentaram aumento significativo na produção de mevalonato, sendo que uma delas produziu 41,5 vezes mais que o controle.

3.3 Integração de transgenes

Assim como para a construção de linhagens KO, CRISPR também se aplica à integração de genes heterólogos. Inicialmente, Dicarlo et al. (2013) demonstraram essa funcionalidade por meio da integração do cassete KanMX no loco *CAN1* com eficiência próximo a 100%. A partir desta premissa, Ryan et al. (2014) testaram a eficiência da integração do gene de resistência à nourseotricina (*Nat^R*) precedida pela montagem *in vivo* de três fragmentos sobrepostos por homologies de 50 pb – promotor, sequência codificadora e terminador. Como resultado, obtiveram 85% de montagem e integração no loco *URA3* para a linhagem S288C diploide e 70% para o poliploide ATCC4124.

Além disso, Ryan et al. (2014) associaram a integração mediada por CRISPR à evolução dirigida de proteínas baseada em *error-prone PCR* para selecionar polipeptídeos com melhor atividade de transporte de celodextrinas. Para isso, escolheram a via derivada do fungo celulolítico *Neurospora crassa*, composta por uma β -glicosidase intracelular (*gh1-1*) e um transportador (*cdt-1*). Em combinação, estes dois genes resultam em linhagens capazes de consumir o dissacarídeo celobiose (Galazka et al., 2010).

Células contendo o gene *gh1-1* previamente integrado no loco *LYP1* (*lyp1::PTDH3-gh1-1-TCYC1*) foram cotransformadas com cinco microgramas do produto de PCR do promotor TDH3; terminador CYC1; biblioteca de alelos mutantes do transportador *gh1-1*; e plasmídeo expressando a Cas9 e o gRNA para o loco *URA3* (P05, Tabela 1). Transformantes selecionados em placas contendo G418 foram, em seguida, testados quanto à capacidade de crescerem em meio sintético contendo celobiose como única fonte de carbono. Como principal resultado desta abordagem, os autores selecionaram células expressando versões do transportador *cdt-1* com maior eficiência na capacidade de utilizar celobiose – aproximadamente duas vezes.

3.4 Integração de vias metabólicas

Devido à instabilidade de plasmídeos episomais, a estratégia de integração é normalmente utilizada para construção de leveduras industriais. Além dos exemplos apresentados, também é possível utilizar CRISPR para promover a integração de múltiplos genes simultaneamente, incluindo vias metabólicas complexas. No entanto, embora *S. cerevisiae* possua excelente habilidade de realizar recombinação homóloga, a eficiência de integração é inversamente proporcional ao tamanho do DNA que se pretende introduzir (Oldenburg et al., 1997).

Neste contexto, ao menos três trabalhos demonstraram o uso de CRISPR como alternativa para integração de grandes construções de DNA incluindo: (i) seis genes da via de produção de ácido mucônico distribuídos em seis fragmentos (Horwitz et al., 2015), (ii) montagem *in vivo* e integração de seis genes do complexo piruvato desidrogenase em um único loco (Mans et al., 2015) e (iii) três genes da via de produção de carotenoides em três locos independentes (Ronda et al., 2015).

Além destes exemplos, Shi et al. (2016) utilizaram CRISPR para integrar múltiplas cópias de ambas as vias de utilização da xilose e da produção do BDO ((R,R)-butano-2,3-diol), que juntas somam 24 mil pb. Para promover múltiplas inserções, os autores acoplaram a habilidade de CRISPR em promover a quebra de fita dupla no DNA – com consequente aumento da frequência de recombinação no sítio de clivagem – à alta ocorrência de sequência Ty-delta no genoma de *S. cerevisiae*. Para isso, os autores utilizaram o vetor previamente desenvolvido por Bao et al. (2015) contendo o gRNA adequado para a edição de sequências Ty-delta (P06, Tabela 1). Este vetor possui ainda uma versão otimizada da Cas9 (iCas9) e o gene *URA3* como marcador de seleção, cuja expressão é dirigida por um promotor truncado (*URA3-d*) que faz com que cada célula mantenha mais cópias do vetor do que o esperado para a origem de replicação 2 μ (*ultra-high copy*) (Loison et al., 1989).

Células da linhagem haploide CEN.PK2-1C foram cotransformadas com 300 ng do vetor contendo todos os elementos para edição via CRISPR e 900 ng do DNA doador – genes das vias de consumo de xilose e produção de BDO. Transformantes foram selecionados em meio seletivo sem uracila e foi constatada a integração de até 18 cópias no genoma de

S. cerevisiae com eficiência próxima a 60%. Em seguida, foi confirmada a capacidade destas linhagens de produzirem BDO a partir de xilose (Shi et al., 2016).

4. Fungos filamentosos

4.1 Introdução

Em comparação às leveduras, fungos filamentosos são mais difíceis de ser geneticamente manipulados de forma precisa. Isto por serem multicelulares, apresentarem diferenciação celular e possuírem células envolvidas por uma parede celular espessa composta por quitina. Além disso, possuem um mecanismo de reparo dominante baseado em NHEJ, o que resulta em inserções de mutações aleatórias na maioria das progênes transformadas. Embora seja possível manipular algumas espécies utilizando conjuntos de ferramentas de biologia molecular como plasmídeos, marcadores de seleção e regiões promotoras específicas (Jiang et al., 2013), para muitas espécies industrialmente relevantes técnicas efetivas ainda estão em fase inicial de desenvolvimento.

Neste contexto, diferentes grupos de pesquisa demonstraram resultados importantes que apontam CRISPR como uma alternativa promissora para mitigar os desafios de edição de genomas de fungos filamentosos. Eles incluem esforços para a manipulação genética de várias espécies representantes do grupo dos ascomicetos como: *Trichoderma reesei* (*Hypocrea jecorina*) (Liu et al., 2015), *Aspergillus spp* (Katayama et al., 2015; Nødvig et al., 2015; Zhang et al., 2016) e *Neurospora crassa* (Matsu-ura et al., 2015).

4.2 Exemplos

Liu et al. (2015) foram os primeiros a demonstrar a funcionalidade de CRISPR no fungo *T. reesei*, espécie amplamente utilizada na produção comercial de enzimas lignocelulolíticas e potencial plataforma para expressão de proteínas heterólogas, devido à sua habilidade de sintetizar e secretar em altas quantidades (Punt et al., 2002). Neste estudo, os autores elaboraram uma estratégia que permite a inativação gênica (*knockout*) ou introdução de novos genes (*knock-in*) por recombinação homóloga em posições precisas do genoma. Além disso, o sistema desenvolvido pode ser usado para gerar modificações simultâneas pela cotransformação de gRNAs e de DNAs doadores para diferentes alvos.

Inicialmente, os autores integram aleatoriamente no genoma de duas linhagens de *T. reesei* uma versão da Cas9 com códons otimizados (*toCAS9*). As linhagens Qm6a e Rut-C30 foram respectivamente transformadas com as construções Ppdc-*toCas9*-Tpdc e Pcbh1-*toCas9*-Tpdc, sendo Ppdc/Tpdc o promotor constitutivo/terminador de uma piruvato descarboxilase e Pcbh1 o promotor do gene celobiohidrolase I, o qual é reprimido por glicose e induzido por oligossacarídeos ou celulose. Em seguida, testes de crescimento, esporulação e da atividade de celulasas foram realizados com transformantes positivos, sendo selecionados aqueles que não apresentaram diferenças quando comparados aos parentais Qm6a e Rut-C30.

Após integração estável da *toCas9*, construiu-se uma linhagem auxotrófica para uracila (*knockout ura5*) a partir da linhagem previamente transformada com a construção Pcbh1-*toCas9*-Tpdc. Uma vez que a expressão de gRNAs demanda, tipicamente, o uso de promotores da RNA polimerase III, os autores optaram por transformar protoplastos com

gRNAs transcritos *in vitro* (vide cap. 7), devido à ausência de caracterização de tais promotores nesta espécie. Após seleção em meio definido contendo 5-Fluoroorotic acid (5-FOA) – análogo tóxico da uracila –, os autores selecionaram uma linhagem *ura5* para testes de recombinação em um ou em múltiplos locos.

Para os testes de recombinação homóloga utilizou-se como marcador de seleção o gene *poura5* derivado de *Penicillium oxalicum*, sendo este controlado pelo promotor *gpdA* e terminador *trpC* (*Pgpda-poura5-Ttrpc*). Inicialmente, protoplastos foram cotransformados com o cassete *Pgpda-poura5-Ttrpc* e com transcritos do gRNA para edição do loco *lae1*, que codifica para uma metil-transferase que controla a expressão de celulases. Após seleção em meio sintético sem uracila, mais de 90% dos transformantes apresentam integração correta do *Pgpda-poura5-Ttrpc* no loco *lae1*, utilizando sequências homólogas de 200 a 1.000 pb.

Aplicando esta mesma estratégia de integração do cassete *Pgpda-poura5-Ttrpc* em uma linhagem auxotrófica, seguida da seleção em meio definido sem uracila, os autores avaliaram a eficiência na edição simultânea de dois e três locos. Para isto, protoplastos foram cotransformados com transcritos de gRNA e DNAs doadores (*Pgpda-poura5-Ttrpc*) contendo homologia adequada para os locos *vib1* e *lae1* ou *vib1*, *lae1* e *clr2*. Após seleção, os autores obtiveram aproximadamente 45% e 4,2% de eficiência para a edição de dois e três locos, respectivamente.

Em outros trabalhos, autores elaboraram estratégias que exploram principalmente a frequência de mutações pontuais geradas durante o reparo via NHEJ para gerar linhagens *knockout*. Em sua maioria, pequenas deleções ou inserções que geram mutações de *frameshift* (Katayama et al., 2015; Nødvig et al., 2015). Como exemplo, Nødvig et al. (2015) desenvolveram uma abordagem de uso simples e versátil, com eficiência de edição satisfatória para ao menos seis espécies do gênero *Aspergillus* (*A. nidulans*, *A. aculeatus*, *A. niger*, *A. carbonarius*, *A. luchuensis* e *A. brasiliensis*).

Esta abordagem consiste em um único plasmídeo contendo os genes que codificam para os dois componentes de CRISPR: Cas9 e gRNA. Ao contrário da estratégia apresentada por Liu et al. (2015), nesta a seleção encontra-se no vetor e quatro versões com diferentes marcadores foram desenvolvidas: *AFUM_pyrG*, *AN_argB*, *bleR* ou *hygR*. Além disso, utilizou-se a sequência AMA1 como origem de replicação, por ser funcional em diferentes fungos (Gems et al., 1991; Ruiz-Díez, 2002).

Ambos os componentes de CRISPR foram modificados. Os códons do gene *Cas9* derivado de *Streptococcus pyogenes* foram otimizados para expressão em *A. niger*. Na extremidade 3' foi adicionada uma sequência codificante para o sinal de localização nuclear SV40. *Cas9-SV40* foram então fusionados ao promotor e terminador *tef1*, ambos de *A. nidulans*. Para a transcrição de gRNAs, os autores optaram por utilizar o promotor constitutivo *gpdA* e o terminador *trpC*, devido à ausência de elementos regulatórios de RNA Polimerase III bem caracterizados também em *Aspergillus*. No entanto, o gRNA transcrito é flanqueado por duas ribozimas – 5' *hammerhead* (HH) e 3' *Hepatitis Delta Virus* (HDV) – que promovem a clivagem do RNA transcrito liberando o gRNA (crRNA + tracrRNA) para posterior interação com a Cas9 (Gao and Zhao, 2014; Jacobs et al., 2014).

Como prova de conceito, os autores construíram linhagens *knockout* para três alelos cuja perda de função é fácil de se monitorar: *yA* gera conídios amarelos em vez de verdes, como observado em linhagens selvagens; *knockout* para o gene *albA* possuem hifas albinas, enquanto a edição do gene *pyrG* gera linhagens resistentes ao 5-FOA.

Outro exemplo de aplicação de CRISPR inclui a edição genômica do fungo *N. crassa*, demonstrada por Matsu-ura et al. (2015). *N. crassa* é considerado organismo-modelo entre os fungos filamentosos e já foi utilizado para elucidar mecanismos moleculares fundamentais como a hipótese “um gene, uma enzima”, ritmo circadiano, entre outros (Roche et al., 2014). No entanto, estudos genéticos baseados em recombinação homóloga nesta espécie só avançaram significativamente a partir da descoberta de linhagens de *N. crassa* com mecanismos de recombinação altamente eficientes, devido à ausência de função dos genes *mus-51* (NCU08290) ou *mus-52* (NCU00077), os quais participam do reparo via NHEJ (Ninomiya et al., 2004).

Para demonstrar a eficiência na substituição de genes mediada por CRISPR, Matsu-ura et al. (2015) utilizaram uma Cas9 contendo o sinal de localização nuclear SV40 e expressão dirigida pelo promotor e terminador *trpC*, ambos derivados de *A. nidulans*. Para expressão do gRNA, utilizaram o promotor SNR52 (snoRNA) e a região 3' do gene SUP4 (tRNA) derivados de *S. cerevisiae* (Harismendy et al., 2003).

Como prova de conceito, os autores optaram por superexpressar celulases a partir da substituição da região promotora do gene *clr-2* pelo promotor constitutivo da β -tubulina (NCU04054). CLR-2 é um fator de transcrição que regula um grande número de celulases em *N. crassa* (Coradetti et al., 2012). Para isto, a linhagem selvagem “74-OR23-1V A” foi cotransformada com cinco microgramas de cada um dos três plasmídeos contendo: (i) a Cas9; (ii) o gRNA e (iii) o DNA doador (5' *clr-2*-P β -tubulin-*clr-2*-*bar*-3' *clr-2* – sendo *bar* o gene que confere resistência ao glufosinato – Ignite®). Aproximadamente 150 transformantes foram obtidos após seleção, e em um dos transformantes avaliados, contendo o gene *clr-2* regulado pelo promotor da β -tubulina, os autores observaram um aumento aproximado de duzentas vezes na expressão de celulases quando comparado à linhagem selvagem.

Neste trabalho, os autores demonstraram o uso eficiente de CRISPR para substituição de genes no fungo *N. crassa*. O resultado mais importante é que, ao contrário dos protocolos já bem estabelecidos para edição genômica em *N. crassa*, esta abordagem não requer que os estudos sejam feitos em mutantes *mus-51* ou *mus-52*, o que possibilita o uso desta tecnologia em isolados naturais da espécie.

5. Plantas

5.1 Introdução

Cientistas buscam há muito tempo uma forma de alterar e melhorar certas características das plantas visando o aumento da produção e resistência a estresses bióticos e abióticos. Até o advento das ferramentas de edição genômica, o desenvolvimento de novas variedades e melhoramento dos cultivares existentes baseava-se no melhoramento tradicional, mutagênese não sítio dirigida, hibridação somática ou métodos de transgenia (Schaeffer & Nakata, 2015).

O desenvolvimento de um conjunto de tecnologias de edição sítio-específica de genomas permitiu o surgimento de uma revolução na área da engenharia de genomas de plantas. Tais tecnologias baseiam-se na utilização de endonucleases sítio-específicas para gerar uma quebra de dupla fita (DSB) no genoma a ser editado, desde que se conheça a sequência do sítio desejado.

Por muitos anos, meganucleases foram utilizadas para gerar DSB, como I-SceI e I-CreI (Li et al., 2015). No entanto, esse método requer a introdução da sequência-alvo no genoma da planta antes do experimento de edição genômica. Por tanto, locos endógenos não podem ser utilizados como alvos (Gimble et al., 2003). Essa limitação foi superada pelo uso das nucleases programáveis do tipo *zinc finger* (ZFNs) e *transcription activator-like effectors* (TALENs).

Atualmente, ZFNs e TALENs estão sendo rapidamente substituídas por endonucleases guiadas por RNA (RGENs, do inglês *RNA-guided endonucleases*) baseada no sistema CRISPR-Cas9, já que esse sistema é mais simples, rápido e barato (Doudna and Charpentier, 2014). Isto se deve ao fato de que CRISPR pode ser customizado através da alteração apenas do componente de RNA, evitando o trabalho intensivo e o consumo de tempo na engenharia de proteínas necessários para a customização dos sistemas TALEN e ZFN.

No início, os estudos com CRISPR foram realizados em células de mamíferos e fungos, demonstrando a grande aplicabilidade desse sistema e também apontando os possíveis problemas como a herdabilidade das alterações geradas e a possibilidade de atividade fora do sítio-alvo. Os cientistas de plantas têm se interessado por esse sistema devido a diversas características, como: facilidade de uso; alta especificidade do local de clivagem e inserção de transgenes; versatilidade na edição de genomas; possibilidade de edição de múltiplos genes ou cópias gênicas ao mesmo tempo e a habilidade de clivagem de locos metilados (revisto em Schaeffer & Nakata, 2015).

Recentemente CRISPR foi utilizada com êxito na edição de genes de diversas plantas, como tabaco, *Arabidopsis*, trigo, milho, sorgo e arroz. Os primeiros estudos foram limitados à prova de conceito e alguns estudos funcionais em que se editam genes que causam alterações fenotípicas facilmente observáveis em plantas-modelos, provando que é possível editar a maioria dos genes endógenos usando a técnica. Também procuraram avaliar especificidades desse sistema quando utilizado em plantas (revisto em Zhou et al., 2014).

Resultados iniciais demonstram que, em plantas, a maioria das quebras de dupla fita geradas por CRISPR são religadas principalmente por NHEJ (Xie et al., 2014). Uma maior eficiência de edição pode ser alcançada através da utilização de variantes de Cas9 com códons otimizados, como foi demonstrado para as variantes de Cas9 otimizadas para arroz (Shan et al., 2013), *nocotiledôneas* (Jiang, et al., 2013) e plantas em geral (Li et al., 2013). CRISPR tem se demonstrado altamente específica, já que nenhum evento de atividade fora do sítio-alvo foi detectado em *Arabidopsis*, tabaco, arroz, laranja e trigo (revisto em Schiml e Puchta, 2016). A maioria dos estudos de CRISPR em plantas também busca demonstrar que as alterações geradas são passadas para as próximas gerações de forma estável.

Nas seções a seguir serão descritos trabalhos que demonstraram a utilidade da técnica para edição do genoma de plantas, através de diversas abordagens do sistema, com vistas a diferentes aplicações como geração de variabilidade genética, nocauteamento de gene, inserção de transgene, deleção de sequências genômicas e regulação da expressão gênica.

5.2 Inserção de transgene (KI) em *Arabidopsis*

Schiml et al. (2014) demonstram que a tecnologia denominada *GT in planta* (de *gene targeting*, Fauser et al., 2012), na qual se faz a inserção de um gene em um sítio-alvo, pode ser usada para inserir um gene de resistência em um loco endógeno através de DSB induzidas por CRISPR em *Arabidopsis thaliana*.

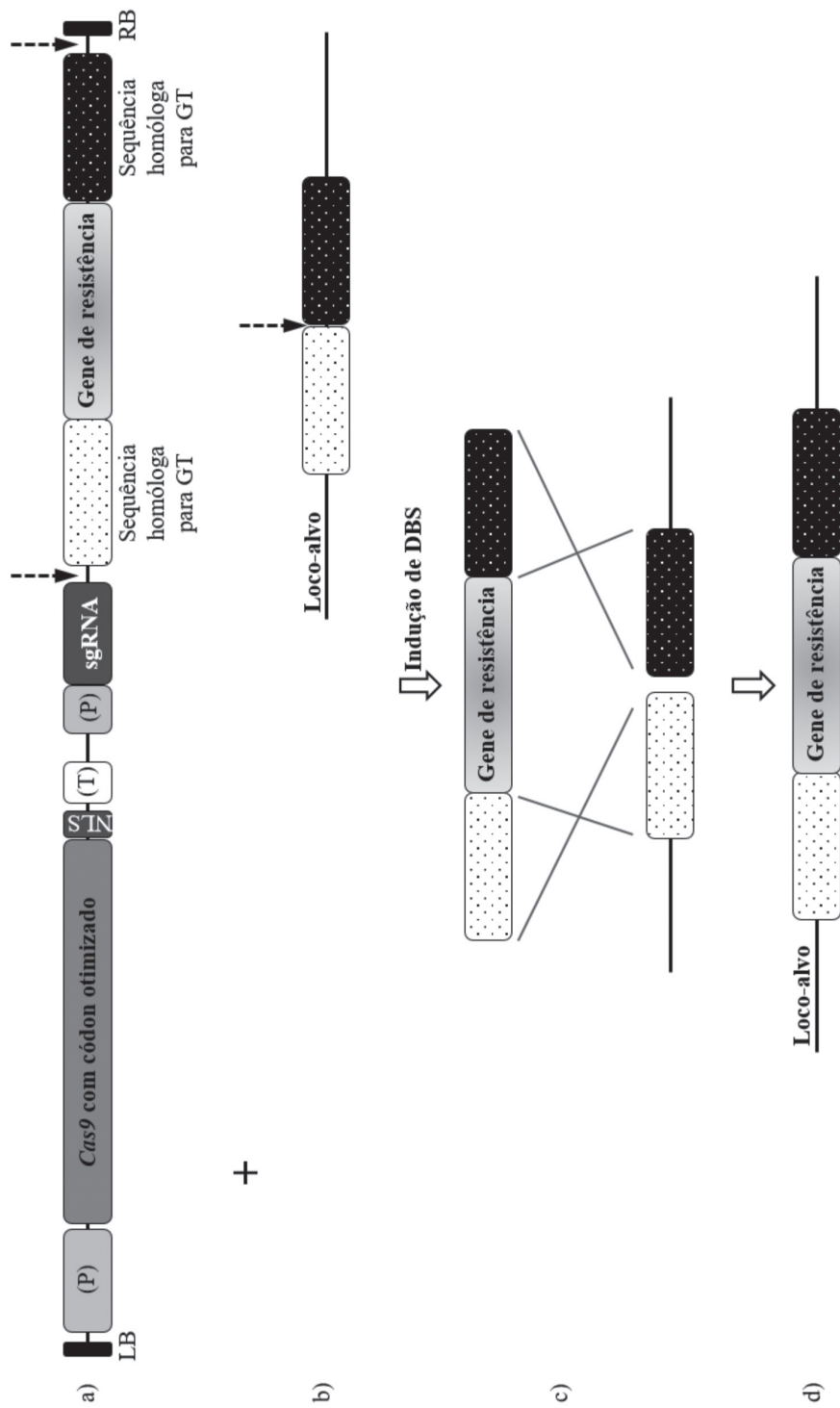


Figura 2. Estratégia para inserção de um gene em um loco-alvo (GT *in planta*) usando CRISPR para gerar quebras de fita dupla (DSB). A) T-DNA contendo as bordas para inserção por *Agrobacterium* (LB e RB); os cassetes *Cas9* e *sgRNA*; e o gene de resistência flanqueado por sequências homólogas ao (B) loco-alvo no genoma (no caso *ADH1*). C) Recombinação homóloga. D) Resultado final da edição. Os sítios para as DSBs estão indicados pelas setas pontilhadas. "P": promotor; "T": terminador; "NLS": sinal de localização nuclear. Imagem: Vieira, AP e Sluys, MA (2016) - adaptado de Schiml et al. (2014).

A técnica *GT in planta* é baseada na presença de uma molécula doadora (DNA doador) e a indução de ao menos três DSBs simultâneas. A molécula doadora do transgene pode ser um T-DNA integrado ao genoma contendo o cassete para GT. Um DSB deve ocorrer no sítio-alvo e duas na molécula doadora, liberando o transgene.

No caso deste estudo, o T-DNA inserido na planta apresenta o sistema de expressão de CRISPR e o cassete GT. O quadro aberto de leitura de Cas9 de *S. pyogenes* utilizado apresenta otimização de códon para *A. thaliana*. A expressão de Cas9 é controlada pelo promotor constitutivo ubiquitina 4-2 de *Petroselinum crispum* (PcUbi4-2). O sgRNA específico para o loco-alvo ADH1 (*alcohol dehydrogenase 1*, AT1G77120) está sob o controle do promotor U6-26 de *A. thaliana* e pode ser facilmente customizado para qualquer sítio de interesse como descrito em Fauser et al. (2014). O cassete GT codifica para o gene de resistência à kanamicina (*nptII*) que é flanqueado por sequências homólogas ao sítio de inserção e dois sítios de reconhecimento para a nuclease Cas9. Esse mesmo sítio de reconhecimento está presente naturalmente no local onde o transgene será inserido. Todas essas partes foram clonadas em um vetor binário para transformação via *Agrobacterium*, que foi utilizado para a transformação da planta por *floral dipping*. A figura 2 representa esquematicamente a estratégia utilizada.

As plantas T1 contendo essa construção foram selecionadas, e a geração T2 foi testada para segregação mendeliana em meio seletivo (contendo kanamicina) para identificar eventos únicos de integração do T-DNA. Os eventos hereditários de *GT in planta* foram identificados através de genotipagem baseada em PCR das gerações T1 a T4. A progênie também foi testada pelo tratamento com álcool alílico, que impede o crescimento das plantas que apresentam o loco ADH1 mutado.

Os resultados obtidos demonstram que as técnicas CRISPR e *in planta* GT podem ser combinadas para a inserção de transgenes de forma sítio-específica no genoma, e os eventos herdáveis podem ser selecionados de forma simples. Essa técnica aumenta o conjunto de opções para engenharia de genomas de plantas e é útil tanto para pesquisa básica quanto para melhoramento, já que pode ser utilizada para adição de genes de interesse em importante cultivares de forma hereditária

5.3 Deleção de sequências genômicas em arroz

Tecnologias que conseguem causar a deleção específica de longas porções genômicas têm o potencial de facilitar análises funcionais de agrupamentos gênicos e elementos genéticos não codificantes. Essas tecnologias também podem ter implicações práticas no melhoramento de plantas através da deleção de locos não desejados.

Zhou et al. (2014) demonstraram que é possível utilizar CRISPR para deletar grandes sequências genômicas em arroz. A estratégia empregada foi a construção de T-DNAs que foram introduzidos em calos derivados de embriões de arroz imaturos através de transfecção por agrobactéria.

Para a construção do T-DNA, eles desenvolveram um conjunto de vetores Cas9/sgRNA baseados no sistema Gateway que são altamente eficientes e fáceis de utilizar. O primeiro vetor contém as bordas LB e LR utilizadas pela agrobactéria, a sequência do cassete SpCas9 com otimização do códon para arroz controlado pelo promotor *maize ubiquitin* (pUbi), gene de resistência à higromicina sob o controle do promotor 35S e sequências de recombinação Gateway. O segundo vetor, vetor intermediário, carrega o cassete sgRNA flanqueado pelos

sítios *att* de Gateway. Isso permite a transferência eficiente do sgRNA desejado para o vetor final.

Primeiramente, eles testaram a eficiência de duas estratégias diferentes de vetores intermediários: 1) vetor contendo tracrRNA e crRNAs separados (dgRNA) e controlados independentemente por promotores U6, com o objetivo de produzir múltiplos crRNAs sob controle de promotores únicos; 2) vetor contendo dois sgRNAs (+85 e +48) também separados e controlados independentemente por promotores U6, com o objetivo de testar qual tamanho de sgRNA é mais eficiente, já que foi demonstrado que sgRNA mais longo (+85) é mais eficiente para mutagêneses em células humanas em cultura (Hsu et al., 2013). Para isso, cada tamanho de sgRNA foi direcionado para diferentes sequências-alvos na região codificante do gene SWEET13.

Os eventos de edição genômica foram detectados através de PCR sobre DNA genômico seguida por ensaio de Endonuclease I T7 (T7E1) nas gerações T0, T1 e T2. Os resultados demonstraram que a construção contendo dgRNAs induz uma taxa muito baixa de mutações, sgRNA (+48) não induz mutações e o sgRNA (+85) produz mutações em altas taxas. Também foi demonstrado que as mutações são estavelmente herdadas.

Assim, os pesquisadores escolheram o sistema Cas9/sgRNA (+85) para demonstrar a prova de conceito e de eficiência de CRISPR na deleção de segmentos cromossômicos. Para tal fim, eles utilizaram como vetor intermediário uma construção contendo dois sgRNAs (+85) separados e controlados independentemente por promotores U6. Esse vetor também foi modificado para conter um sítio de restrição *Hind*III, de modo que outro par de cassetes sgRNAs (+85) pudesse ser inserido, permitindo a expressão de quatro diferentes sgRNAs.

Eles testaram a deleção de duas regiões diferentes do genoma: um agrupamento de cinco genes envolvidos na produção de diterpenoides, que são um grupo de fitoalexinas com atividade antimicrobiana, com tamanho de 170 kb no cromossomo 4; e um agrupamento de dez genes também envolvidos na produção de diterpenoides, de aproximadamente 245 kb no cromossomo 2. Para cada região, foram escolhidos dois genes-alvos (um em cada ponta). Para cada gene, duas sequências-alvos foram definidas, de modo que para cada região foram construídos quatro sgRNAs, de forma a permitir testar a deleção por dois diferentes pares e também a funcionalidade de um *multiplex* de quatro sgRNAs.

As construções obtidas foram transferidas para protoplastos de arroz, que foram analisados após 48h das transfecção. A análise foi feita através de PCR sobre o DNA genômico, seguido de sequenciamento dos amplicons. Os resultados mostraram que a deleção cromossômica desejada foi obtida para ambas as regiões (170 kb e 245 kb) com quaisquer das combinações de dois ou quatro sgRNAs, provando que o sistema é altamente eficiente. Além disso, os autores também comprovaram que essas deleções são estáveis através da transformação das construções em calos de arroz por agrobactéria, regenerados em meios seletivos contendo higromicina. As plantas regenerantes foram analisadas através de PCR e sequenciamento para verificar a deleção.

5.4 Geração de variabilidade genética livre de DNAs exógenos

Edição de genomas de plantas sem a introdução de DNA exógeno pode aliviar as preocupações regulatórias relacionadas a organismos geneticamente modificados (GMOs) (Wook Woo et al., 2015). As pequenas inserções ou deleções (*indels*) ou substituições geradas

por CRISPR nos sítios-alvos cromossômicos são indistinguíveis de variações genéticas que ocorrem de forma natural. Em abril de 2016, o departamento de agricultura dos Estados Unidos da América (USDA) decidiu não regular um cogumelo com genoma editado por CRISPR, apesar de ser um GMO. Esse cogumelo pode, então, ser cultivado e vendido sem passar pelo processo de regulamentação da agência, tornando-se o primeiro organismo editado via CRISPR a receber um sinal verde do governo dos EUA (Waltz, 2016).

A expressão do “cassete CRISPR” (um vetor contendo a sequência para a proteína Cas9 e a sequência sgRNA) pode ser feita de forma transiente, na qual o vetor de expressão é inserido na célula da planta mas não é integrado ao genoma, ou de forma permanente, na qual, através de biobalística ou infecção por *Agrobacterium tumefaciens*, o vetor de expressão é integrado ao genoma da planta. Quando a expressão permanente é utilizada, a planta resultante terá o DNA editado no sítio-alvo, porém, também apresentará um DNA exógeno, que é composto pela integração aleatória do T-DNA no genoma. Como a edição do genoma ocorre logo após a expressão do “cassete CRISPR”, o T-DNA (que contém Cas9/sgRNA) pode ser removido do genoma da planta, já que não há necessidade de expressar o sistema continuamente. Esse DNA exógeno pode ser eliminado através de cruzamentos genéticos segregando o T-DNA. No entanto, essa remoção não é factível em espécies como uva, batata e banana, cuja propagação é vegetativa. A expressão transiente pode ser utilizada, no entanto, esses plasmídeos são degradados por endonucleases endógenas, resultando em pequenos fragmentos de DNA que podem ser inseridos tanto no sítio-alvo quanto em outras posições, gerando assim plantas que também contêm DNA exógeno.

Para gerar plantas com os genomas editados, livres de DNA exógeno, que talvez possam ser isentadas das leis regulatórias para GMOs, Wook Woo et al. (2015) utilizaram uma abordagem *in vitro*, livre de DNA, da tecnologia CRISPR, que foi previamente demonstrada em cultura de células humanas (Kim et al., 2014). Nesta proposta, os autores utilizaram a ribonucleoproteína (RNP) formada pela interação da proteína Cas9 com o sgRNA pré-montada *in vitro* (denominada RGN RNP) para inserir em protoplastos de diversas plantas e regenerar plantas inteiras editadas.

Primeiramente, para formar a RNP, a proteína Cas9 purificada é misturada com o sgRNA obtido através da transcrição *in vitro*. Então diferentes RNPs obtidas foram incubadas com protoplastos de diferentes plantas na presença de polietileno glicol (PEG). Os genes-alvos são: *PHYB* (*A. thaliana*), *AOC* (*N. attenuata*) e *P450* e *DWD1* (*O. sativa*). Para avaliar a frequência de mutações, foram feitos ensaio de Endonuclease I T7 (T7E1) e sequenciamento por *deep sequencing*. *Indels* foram identificados nas posições esperadas em todos os casos, com frequências que variam de 8,4% a 44%.

Os pesquisadores também testaram essa técnica de RGN RNP em protoplastos de alface para o nocauteamento do gene BRASSINOSTEROID INSENSITIVE 2 (*BIN 2*) homólogo ao de *A. thaliana*, que codifica um regulador negativo da via de sinalização de brassinosteróide. O mesmo RGN RNP utilizado na edição foi usado na análise de RFLP para a genotipagem dos calos de alface obtidos. Esse ensaio mostrou que 5,7% dos calos continham mutações monoalélicas e 40% continham mutações bialélicas no sítio-alvo. Plantas inteiras foram regeneradas a partir desses calos e sementes de mutantes bialélicos homozigóticos foram germinadas e apresentaram a mutação, comprovando a herdabilidade. Contudo, estudos para avaliar se o mutante apresenta o fenótipo de aumento na sinalização de brassinosteróide previsto ainda não foram publicados.

5.5 Regulação da expressão gênica *in planta*

A modulação da expressão gênica de maneira direcionada é uma abordagem poderosa para acelerar o estudo funcional de forma eficiente e precisa, além de permitir novas estratégias para o desenvolvimento da biotecnologia agrícola. Piatek e colaboradores (2015) apresentam a primeira evidência de que CRISPR pode ser utilizada para controlar a expressão gênica direcionada em plantas. Tal modificação se baseia na proteína Cas9 inativada cataliticamente (dCas9) que, combinada com um sgRNA, atua como um módulo sintético de direcionamento para DNA. Bikard et al. (2013) e Qi et al. (2013) demonstraram que a ligação de dCas9 a DNAs-alvos interfere na transcrição gênica.

Sendo assim, neste estudo os autores construíram algumas versões de dCas9 (com códon otimizado para humanos), cada qual fusionada a diferentes proteínas envolvidas na regulação da expressão gênica, permitindo o controle seletivo de genes-alvos *in planta*. Para gerar ativadores de transcrição, a porção C-terminal da proteína dCas9 foi fusionada aos domínios de ativação EDLL (que são fortes ativadores e atuam próximos ou distantes do promotor-alvo) e efetores TAL (que são fatores de transcrição *in planta*). Para gerar repressores de transcrição, a porção C-terminal de dCas9 foi fusionada ao domínio de repressão SRDX (que é um repressor potente e dominante). Todas as construções foram clonadas no vetor de expressão de plantas pK2GW7. Os sgRNAs específicos para cada abordagem sob o controle do promotor U6 foram clonados em outro vetor binário. A figura 3 representa esquematicamente as estratégias utilizadas e os mecanismos de regulação proposto.

Para testar o efeito dos ativadores de transcrição sintéticos dCas9:TAD e dCas9:EDLL em um alvo endógeno, no caso o gene *Phytoene desaturase* (*PDS*), três sgRNAs com diferentes sítios-alvos no promotor de *PDS* foram utilizados. Os dois vetores – um contendo os ativadores de transcrição sintéticos dCas9:TAD ou dCas9:EDLL e outro contendo um dos sgRNAs – foram separadamente transformados em *Agrobacterium* e coentregues em diferentes combinações em folhas *N. benthamiana* via agroinfiltração. Discos das folhas foram coletados após 36-48h da infiltração. A quantidade de transcritos de *PDS* foi analisada por RT-qPCR. Os resultados indicam que tanto dCas9:TAD quanto dCas9:EDLL foram capazes de ativar, de forma significativa, a transcrição do gene-alvo.

O efeito repressor da construção dCas9:SRDX foi testada com *PDS* como gene-alvo. As mesmas estratégias de transformação de *N. benthamiana* e análise de dados foram utilizadas e também demonstraram a efetividade do sistema também na repressão de genes, corroborando que essa plataforma baseada em dCas9 pode ser usada para a regulação direcionada de genes *in planta*.

O desenvolvimento de reguladores de transcrição sintéticos baseados em dCas9 podem ser utilizados em plantas-modelos para facilitar análises funcionais de genes sob diferentes condições nutricionais, hormonais ou ambientais. Uma abordagem interessante seria utilizar linhagens estáveis de plantas contendo os cassetes dCas9:fatores de transcrição e apenas transformar com sgRNA através de vetores virais que não causam sintomas, de forma a permitir a regulação sistêmicas de genes-alvos.

5.6 Próximos passos do uso de CRISPR em plantas

Com a maioria dos trabalhos conceituais acerca de CRISPR, pesquisadores de plantas estão começando a aplicar essa tecnologia de edição genômica no melhoramento

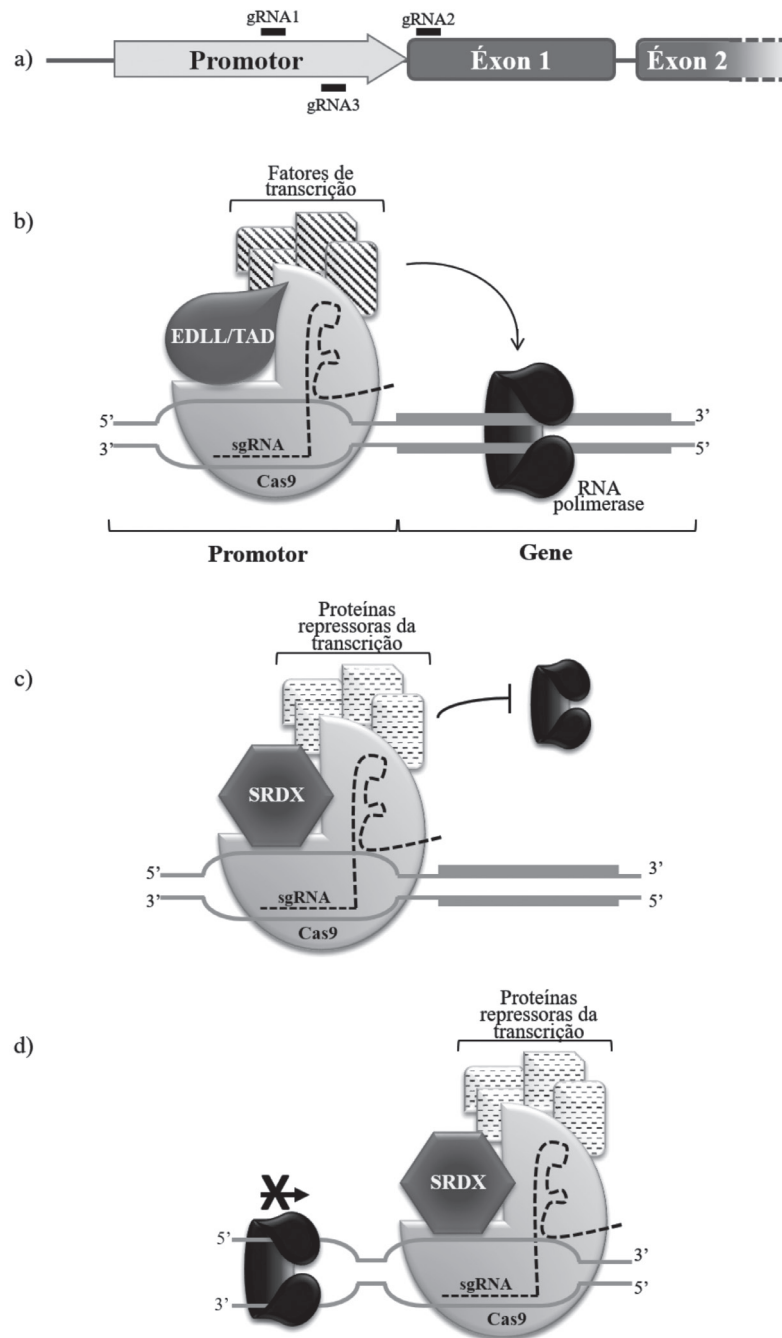


Figura 3. Estratégias para regulação da transcrição gênica através do sistema CRISPR-Cas9. A) Representação esquemática do gene-alvo, para o qual três gRNAs são desenhados tendo como alvo a região promotora e o primeiro éxon. Em todas as situações, dCas9:reguladores de transcrição são direcionados por sgRNAs para o sítio-alvo. Os reguladores atraem fatores de transcrição. B) Representação esquemática do mecanismo de ativação da transcrição gênica. C) e D) Representam mecanismos de repressão da transcrição gênica. Em C), o complexo bloqueia o início da transcrição, enquanto em D) a elongação da transcrição é interrompida. Imagem: Vieira, AP e Sluys, MA (2016) - modificado de Piatek et al. (2015).

de características específicas em cultivares de importância econômica. A Tabela 2 mostra exemplos de usos da técnica na edição de genes para melhoramento.

Antes de estratégias racionais serem projetadas para implementar esta tecnologia na obtenção de grandes variedades de culturas de importância econômica, há uma necessidade de expansão da disponibilidade de vetores para culturas específicas, disponibilidade de genomas sequenciados e protocolos de transformação e regeneração.

Além disso, a tecnologia CRISPR-Cas9 continua evoluindo para além da edição genômica. Por exemplo, pesquisadores foram capazes de reprogramar CRISPR para reconhecer e clivar RNAs específicos em vez dos usuais alvos de DNA (Nelles et al., 2015; O'Connell et al., 2014). Essa habilidade abre novos caminhos na detecção, análise a manipulação de RNAs.

6. Animais

Este capítulo visa apresentar diferentes métodos de CRISPR, em diferentes animais com variadas aplicações, para que o leitor tenha uma visão panorâmica das inúmeras possibilidades que a técnica possui.

6.1 *Drosophila melanogaster* (Chen et al., 2016)

Grande parte da biodiversidade terrestre evoluiu em exposição às variações cíclicas diárias determinadas pelo movimento de rotação da terra. Os comportamentos periódicos diários são parte do chamado ciclo circadiano, e são muito importantes para a manutenção da homeostase. Seus mecanismos regulatórios têm origens profundas na fisiologia e biologia molecular, são compartilhados por grande parte dos organismos, e alterações em seus genes-chave estão relacionadas a doenças e síndromes metabólicas (Takahashi et al., 2008).

Nesse exemplo, CRISPR foi usada para a edição do genoma de *Drosophila melanogaster* visando eliminar o sítio de ligação de um microRNA (mir-276a) na região 3' UTR do gene *timeless* (*tim*) (Chen and Rosbash, 2016). A proteína TIM é parte de um circuito molecular regulatório do relógio circadiano baseado em um *loop* em *feedback*. Nesse circuito, a transcrição e a tradução dos genes participantes é essencial para determinar a periodicidade adequada. MicroRNAs são capazes de modular a tradução de proteínas-alvo através da sua ligação no 3' UTR de mRNAs. A hipótese dos pesquisadores era de que mir-276a estaria envolvido diretamente na modulação do circuito molecular circadiano de *Drosophila* através da sua ligação ao 3' UTR do mRNA do gene *timeless*. Os pesquisadores conseguiram demonstrar que quebras de fita dupla mediadas por Cas9 foram suficientes para gerar alelos com deleções de até 10 pb no sítio de ligação de mir-276a, abolindo o seu papel regulatório, e fazendo com que a proteína TIM acumulasse. De modo importante, isso resultou na alteração da periodicidade do circuito molecular com consequente perda do controle do ciclo de atividade nas moscas geneticamente alteradas. Isso permitiu elucidar um novo mecanismo de modulação desse circuito baseado na atividade de um microRNA.

6.2 *Caenorhabditis elegans* (Paix et al., 2014)

De forma semelhante ao peixe-zebra, o nematoide *C. elegans* tornou-se um importante modelo em genética da biologia do desenvolvimento. Devido ao seu tamanho e acessibilidade visual, todas as etapas morfológicas necessárias ao desenvolvimento de um verme adulto

Tabela 2. Exemplos de edição gênica para melhoria de plantas utilizando CRISPR. Muitas das características alteradas não são economicamente viáveis, no entanto, o estabelecimento de métodos para produzir mutações bialélicas em uma espécie representa um primeiro passo na transição destas tecnologias para o desenvolvimento de novas cultivares.

Espécie	Gene(s)-alvo	Eficiência de edição ^a	Característica alterada	Técnica de CRISPR	Efeito molecular	Técnica de transformação/ transfeção	Referência
<i>Glycine max</i>	<i>GmDD1</i>	~0-14%	Atividade da DNA metiltransferase	Quebra única de fita dupla	Modificações diversas ^b	Biobalística	Jacobs et al., 2015
<i>Nicotiana tabacum</i>	<i>NtPDS</i>	~82%	Folhas estioladas	Quebra única	INDELS causando nocaute	Mediada por <i>Agrobacterium</i>	Gao et al., 2015
	<i>NtPDR6</i>	~88%	Ramificação	de fita dupla			
<i>Nicotiana tabacum</i>	<i>NtPDS</i>	~82%	Folhas estioladas	CRISPR	Deleções e inversões de 1,8kb	Mediada por <i>Agrobacterium</i>	Gao et al., 2015
	<i>NtPDR6</i>	~88%	Ramificação	multiplexing			
<i>Oryza sativa</i>	<i>OsSWEET1a</i>	70%					
	<i>OsSWEET1b</i>	20%	Resistência a crestamento	CRISPR	Nocaute de diversos genes	Mediada por <i>Agrobacterium</i>	Zhou et al., 2015
	<i>OsSWEET13</i>	60% - 100%	bacteriano	multiplexing			
	<i>OsSWEET11</i>	12,5%					
<i>Oryza sativa</i>	Agrupamento <i>OsCPS4/</i>						
	<i>CYP99A2</i>	4,2%	Síntese de diterpenoides	CRISPR multiplexing	Deleção de genes	Mediada por <i>Agrobacterium</i>	Zhou et al., 2016;
<i>Oryza sativa</i>	<i>OsBEL</i>	~2-16%	Sensibilidade a herbicida	CRISPR multiplexing	Nocaute	Mediada por <i>Agrobacterium</i>	Xu et al., 2014
<i>Populus tomentosa</i>	<i>ProPDS</i>	51,7%	Folhas estioladas	Quebra única de fita dupla	Nocaute	Mediada por <i>Agrobacterium</i>	Fan et al., 2015

^aA eficiência de edição foi medida por diferentes métodos em cada estudo.

^bOs eventos de DSB resultaram em diferentes mutações, incluindo deleções, SNPs, inserções e substituições (duas ou mais bases que foram incorporadas depois de um evento de deleção).

Tabela 2. Continuação.

Espécie	Gene(s)-alvo	Eficiência de edição ^a	Característica alterada	Técnica de CRISPR	Efeito molecular	Técnica de transformação/transfecção	Referência
<i>Solanum lycopersicum</i>	<i>SIAGO7</i>	48%	Forma da folha	Quebra única de fita dupla	Nocaute por deleção de fragmento	Mediada por <i>Agrobacterium</i>	Brooks et al., 2014
<i>Solanum tuberosum</i>	<i>StIAA2</i>	50%	Hiponastismo do pestiolo e morfogênese da parte aérea	Quebra única de fita dupla	Nocaute	Mediada por <i>Agrobacterium</i>	Wang et al., 2015
<i>Triticum aestivum</i>	<i>TaMLO-A1</i>	5,6%	Resistência a oídio	Quebra única de fita dupla	Nocaute de alelos	Biobalística	Wang et al., 2014
<i>Zea mays</i>	<i>ZmIPK</i>	13,1	Síntese de ácido fítico	Quebra única de fita dupla	Nocaute	Mediada por <i>Agrobacterium</i>	Liang et al., 2014
<i>Zea mays</i>	<i>ZmHKT1</i>	?	Transporte de íons	CRISPR multiplexing	Nocaute de diversos genes ao mesmo tempo	Mediada por <i>Agrobacterium</i>	Xing et al, 2014

^aA eficiência de edição foi medida por diferentes métodos em cada estudo.

^bOs eventos de DSB resultaram em diferentes mutações, incluindo deleções, SNPs, inserções e substituições (duas ou mais bases que foram incorporadas depois de um evento de deleção).

são conhecidas para cada célula gerada a partir das clivagens iniciais do zigoto. Apesar de ser um organismo permissivo à manipulação genética, a possibilidade da realização de edições precisas em seu genoma sem a necessidade de utilização de marcadores de seleção ainda envolve grande esforço na identificação das linhagens geneticamente modificadas.

Nesse exemplo, os pesquisadores exploraram múltiplas formas de utilizar CRISPR para a edição do genoma de *C. elegans* via reparo por recombinação homóloga mediado por diversas sequências de DNA doadoras (plasmídeos, produtos de PCR e oligonucleotídeos de fita simples-ssODN). No experimento mais simples e bem-sucedido, eles almejavam a edição de diversos locos individualmente por meio da inserção de alterações na sequência genômica-alvo via cotransformação de sgRNAs e Cas9 em conjunto com ssODN. Esses ssODN podem ser rapidamente sintetizados a baixo custo e se mostraram excelentes para introduzir pequenas alterações no genoma, como mutações pontuais, alterações do quadro de leitura, códons de terminação prematuros, e caudas para fusões nas regiões amino- e carboxi-terminal de proteínas-alvo. Para isso, essas modificações desejadas são incluídas no centro do ssODN e flanqueadas por sequências homólogas ao 5' e 3' do sítio-alvo determinado pela escolha do sgRNA. As maiores frequências de reparo exato por recombinação homóloga ocorreram quando regiões de homologia imediatamente adjacentes ao sítio de corte da Cas9 foram utilizadas, permitindo a edição precisa do genoma de *C. elegans* sem a necessidade de utilização de vetores doadores grandes contendo marcadores de seleção.

6.3 *Xenopus laevis* (Wang et al., 2015)

Xenopus laevis são modelos não mamíferos investigados na biologia moderna desde os anos 1950, com importantes contribuições em áreas como bioquímica, biologia molecular e do desenvolvimento. Neste exemplo, CRISPR foi empregada com sucesso para promover edição gênica em *X. laevis* pela primeira vez. Através da coinjeção do mRNA de Cas9 e de sgRNAs específicos, foram obtidos modelos com os genes *ptf1a/p48*, associado à formação do pâncreas, e *tyr* (tirosinase), envolvido na pigmentação, inativados. Para *ptf1a/p48*, apenas um sgRNA foi injetado em embriões de *Xenopus*, e uma taxa de eficiência de edição gênica de 50% foi observada. Para o gene *tyr*, dois sgRNAs foram injetados, um para cada par de genes (*tyra* e *tyrb*). A técnica mostrou-se extremamente eficiente, com taxas de sucesso variando de 87,5 a 100%. Esses resultados demonstraram que CRISPR pode ser utilizada para avaliar características fenotípicas precocemente em embriões G₀ de *X. laevis*.

6.4 *Zebrafish* (Di Donato et al., 2016)

Devido ao tamanho, tempo rápido de geração e transparência dos tecidos, o peixe-zebra (ou peixe paulistinha) tem sido cada vez mais utilizado como um modelo em biologia do desenvolvimento e neurobiologia. Neste contexto, a capacidade de se inativar genes-alvo em tecidos ou até mesmo populações de células específicas, permitindo ao mesmo tempo a sua identificação e isolamento posterior, é extremamente desejável.

Neste exemplo, os pesquisadores obtiveram uma linhagem de peixe transgênica expressando constitutivamente um sgRNA direcionado à inativação do gene da tirosinase, essencial para a síntese de pigmento, além de conter o gene Cas9 fusionado a um repórter GFP via um peptídeo viral do tipo 2A (Cas9:P2A:GFP), ambos sob controle do promotor exógeno *Gal4*. Esse sistema permite a expressão de Cas9 de modo controlado ao se cruzar essa linhagem com uma das muitas disponíveis expressando o fator de transcrição Gal4 sob

controle de promotores célula ou tecido-específicos. Além disso, a construção Cas9:P2A:GFP faz com que durante a tradução do transcrito de fusão, as proteínas Cas9 e GFP sejam separadas pela autoclivagem da sequência peptídica P2A, permitindo a identificação qualitativa e quantitativa da expressão de Cas9 baseada na fluorescência do GFP. Como prova de conceito da ferramenta gerada, o cruzamento dessa linhagem de peixe com uma linhagem *Rx2-Gal4* específica dos primórdios óticos embrionários resultou em peixes com defeitos de pigmentação em células da retina. Conforme o esperado, as mesmas populações celulares afetadas pela falta de pigmentação mostraram-se GFP positivas, o que também permite o seu isolamento por citometria de fluxo para estudos bioquímicos futuros.

6.5 Camundongos (Maddalo et al., 2014)

Camundongos (*Mus musculus*) são provavelmente o modelo biológico mais extensivamente utilizado na pesquisa biomédica. Apesar das facilidades de manipulação genética e o vasto conhecimento acumulado sobre a sua biologia, diversos aspectos de doenças humanas dificilmente são reproduzidos nesse modelo animal. Particularmente, a modelagem de síndromes neoplásicas associadas a mutações somáticas restritas a órgãos específicos é muito limitada nas abordagens clássicas de transgenia.

No exemplo seguinte, os pesquisadores buscaram reproduzir em camundongo uma mutação somática recorrente em tumores pulmonares do tipo Carcinoma de Células Não Pequenas (CNPC, em inglês). O desafio foi recapitular precisamente a mutação recorrente em pacientes humanos, não no organismo todo, mas sim de forma localizada no tecido pulmonar. Para isso, um vetor adenoviral foi utilizado para administrar a proteína Cas9 em conjunto com dois sgRNAs nos pulmões de camundongos via infiltração. Os sgRNAs tinham como alvo íntrons dos genes *Alk* e *Eml4*, frequentemente envolvidos em uma inversão responsável pela formação do oncogene *Eml4-Alk*, comum nos tumores CNPC. Como resultado, os camundongos tratados desenvolveram tumores pulmonares do tipo CNPC com patologia extremamente similar à dos tumores humanos. Os tumores continham a presença do oncogene *Eml4-Alk*, cuja origem pôde ser relacionada à indução de quebras no DNA pela proteína Cas9 associada aos sgRNAs utilizados. Essa estratégia permitiu a obtenção do primeiro modelo de CNPC em camundongos. Também permitiu, pela primeira vez, a experimentação de terapias com inibidores específicos, ensaios de sensibilidade e de resistência a diferentes drogas experimentais para o tratamento de CNPC em humanos (Garcia-Bloj et al., 2015).

6.6 Ratos (Bakondi et al., 2016)

Um alelo mutante do gene da rodopsina (*Rho*) é considerado uma das causas mais associadas à retinite pigmentosa (RP), uma doença genética com caráter autossômico dominante. Ratos contendo o alelo de camundongos *RhoS334* desenvolvem similaridades fenotípicas com a RP humana, sendo um modelo-chave para o estudo desta patologia em humanos.

Neste exemplo, CRISPR foi utilizada para inativar *in vivo* a versão mutante do alelo *Rho*. Um sgRNA específico para o alelo mutante foi projetado diferindo do alelo selvagem por possuir uma substituição de G por A na posição 10 do sgRNA e uma base G (TGG) ao invés de C (TGC) formando o motivo PAM, essencial para a atividade da Cas9. Os animais receberam uma injeção sub-retinal de um plasmídeo contendo o sgRNA e a sequência da Cas9, seguida de eletroporação *in vivo*. Posteriormente, as células da retina foram isoladas e genotipadas, com edição gênica presente em até 36% especificamente nos alelos mutantes.

Em comparação com o grupo controle, que geralmente apresenta degeneração total da retina 28 dias após o nascimento, foi observada uma melhora significativa no número de camadas de fotorreceptores, prevenção da degeneração de retina e melhora na função visual dos animais editados *in vivo* via CRISPR. Este foi o primeiro exemplo de correção *in vivo* de uma doença autossômica dominante utilizando a técnica, abrindo portas para o emprego desta tecnologia em aplicações semelhantes em humanos.

6.7 Cabras (Ni et al., 2014)

Caprinos posicionam-se como uma espécie estratégica para a manipulação genética, tanto para o estudo de doenças humanas quanto para a produção de proteínas terapêuticas, tendo como maior exemplo a aprovação para o uso humano de uma antitrombina humana recombinante purificada a partir do leite de cabras transgênicas (Kling, 2009).

Estes pequenos ruminantes possuem um intervalo entre gerações de cinco meses, o que limita a aplicação de múltiplas modificações genéticas utilizando-se técnicas tradicionais pelo tempo e alto custo envolvido. Entretanto, o advento das nucleases para edição de DNA está mudando este panorama, permitindo que diversos locos sejam editados concomitantemente. Neste exemplo, CRISPR foi empregada com sucesso para o nocaute simultâneo de até quatro genes codificadores para proteínas envolvidas na regulação do crescimento muscular (miostatina – *MSTN*), doença neurodegenerativa (proteína do prion – *PrP*), modelo para o estudo de doenças cardíacas (nucleoporina – *NUP*) e alergia ao leite (beta-lactoglobulina – *BLG*) em fibroblastos caprinos. Surpreendentemente, foram obtidas colônias com nocautes bialélicos nos quatro genes-alvo em apenas uma cotransfecção de quatro plasmídeos contendo os sgRNAs específicos para cada gene em conjunto com o gene para a Cas9. Duas colônias contendo mutações bialélicas para *MSTN* foram utilizadas na clonagem por transferência nuclear de células somáticas, resultando em três clones nascidos, dos quais dois sobreviveram. Todos os animais foram genotipados por RFLP e confirmaram as mesmas mutações oriundas das células doadoras de núcleo para a clonagem. Possíveis sítios de clivagens fora de *MSTN* (*off-targets*) foram investigados nos animais clonados, mas nenhuma edição gênica além da esperada foi observada. Em adição, não foi detectada expressão de *MSTN* no tecido muscular dos clones caprinos. Além da edição de DNA em células associada à clonagem para a obtenção de cabras geneticamente modificadas, CRISPR foi empregada com sucesso para o nocauteamento de genes também por meio da microinjeção de zigotos (Wang et al., 2015). Esses resultados demonstram a versatilidade do sistema para animais de produção e abrem portas para modificações mais complexas, que certamente estão por vir.

6.8 Bovinos (Heo et al., 2015)

Bovinos são espécies de destacado interesse agropecuário, principalmente pela elevada produção de carne e leite. A produção de gado geneticamente modificado pode atender às demandas tanto da indústria alimentícia quanto farmacêutica, mas atualmente ainda esbarra na relativa ineficiência das técnicas para a modificação genética, como *gene-targeting* em células somáticas, clonagem e microinjeção de zigotos.

Buscando encontrar alternativas eficientes para a manipulação do genoma bovino, pesquisadores utilizaram CRISPR para editar o genoma de células-tronco pluripotentes induzidas (iPSCs) e para introduzir o gene da proteína verde fluorescente (GFP) precisamente no genoma tanto de iPSCs quanto de embriões bovinos. iPSCs foram cotransfectadas com

um vetor com Cas9/sgRNA além de um vetor doador contendo GFP e o gene de resistência ao antibiótico puromicina. A técnica foi altamente eficiente, com 100% de integração no referido loco e expressão de GFP confirmada por citometria de fluxo nas colônias resistentes ao antibiótico. Para verificar a eficiência do sistema na produção de embriões transgênicos, o mRNA de Cas9 e o gRNA foram coinjetados em zigotos juntamente com o vetor doador para a inserção por recombinação homóloga de GFP no íntron 1 do gene *NANOG*. Esta estratégia obteve sucesso, com até 13,9% de blastocistos expressando GFP.

Apesar do sucesso das manipulações genéticas *in vitro*, ainda não foram descritos bovinos nascidos vivos com o genoma editado por CRISPR. Entretanto, o já reportado sucesso de outras nucleases de edição de DNA, como TALENs e ZFNs (Tan et al., 2016), aliado aos expressivos resultados obtidos *in vitro* (Heo et al., 2015; Jeong et al., 2015), sugerem que o uso de CRISPR para produzir bovinos geneticamente modificados é apenas uma questão tempo.

6.9 Cães (Zou et al., 2015)

Cães são espécies domésticas que, além de animais de companhia, são extremamente úteis como modelos para pesquisas biomédicas. Em razão de suas características reprodutivas peculiares, modificações genéticas nesta espécie são raras.

Neste exemplo, que é o único trabalho publicado até o presente momento empregando CRISPR em cães, os pesquisadores editaram o gene *MSTN*, responsável pela produção da proteína miostatina, reguladora negativa do crescimento muscular. Inicialmente foram realizados alotransplantes, em que embriões microinjetados com o mRNA de Cas9 e um sgRNA direcionado à sequência codificante de *MSTN* foram transferidos para fêmeas receptoras, mas não resultaram em prenhez viáveis. Em uma segunda estratégia, embriões coletados em um dos ovidutos foram microinjetados e autotransplantados no oviduto contralateral da mesma fêmea. Esta estratégia obteve sucesso e, dos 27 cães nascidos, dois apresentaram edição no gene *MSTN*, sendo um bialélico e o outro um mosaico. O animal com mutação bialélica não apresentou expressão de miostatina, e análises de ressonância magnética confirmaram o padrão esperado de dupla musculatura. Amostras de sêmen de um dos animais apresentaram as mutações esperadas em *MSTN*, confirmando que as modificações genéticas obtidas com o emprego de CRISPR em cães podem ser transmitidas para a linhagem germinativa.

6.10 Suínos (Yang et al., 2015)

Anualmente no Brasil cerca de 20.000 pessoas encontram-se na lista de espera para um transplante de órgãos (ABTO, 2015). Em proporções semelhantes, a alta demanda de órgãos para transplante é um problema global. O xenotransplante (transplante de órgãos animais em humanos) é visto como uma promissora solução, no entanto, essa tecnologia é limitada ainda devido às fortes reações imunológicas cruzadas e ao risco de infecção das células do paciente receptor por vírus endógenos presentes no genoma (ou endoviroma) do animal doador.

No exemplo seguinte, os pesquisadores se utilizaram de CRISPR para inativar todas as 62 cópias retrovirais endógenas do genoma do porco (*Sus scrofa domesticus*) em um único passo. Dado o alto grau de similaridade entre as cópias, a utilização de apenas dois sgRNAs mostrou-se suficiente para alvejar todo o endoviroma. Em seguida, os pesquisadores utilizaram um sistema de transfecção baseado no transposon *PiggyBac* para introduzir os

sgRNAs e o gene da proteína Cas9 sob o controle de um promotor induzível por doxiciclina em células PK15. Na presença contínua de doxiciclina por 17 dias, um percentual significativo de populações clonais das células sofreu a inativação completa do seu endoviroma por mutações induzidas pela Cas9. A estratégia não induziu nenhuma mutação secundária detectável, e as células resultantes mostraram-se virtualmente incapazes de infectar células humanas com vírus endógenos de porco em cocultura. Os cientistas também reportaram a obtenção de clones de porcos obtidos através da transferência nuclear de células somáticas, gerando a primeira linhagem de animais com inativação completa do endoviroma. Essa estratégia foi não somente um exemplo pioneiro de que CRISPR pode ser utilizada para edição simultânea de locos repetitivos, mas também um passo significativo para tornar xenotransplantes uma alternativa viável para milhares de pessoas na fila de espera para receber doações de órgãos (Servik, 2015).

6.11 Macacos (Chen et al., 2015; Kang et al., 2015)

Apesar de serem amplamente utilizados como modelos para o estudo de doenças, camundongos nem sempre reproduzem fidedignamente todas as características de patologias observadas em humanos. Nesse contexto, métodos eficientes para a obtenção de modificações genéticas em macacos são altamente desejáveis para pesquisas no campo biomédico, por estes apresentarem alta semelhança genética e fisiológica com a espécie humana.

Em um primeiro estudo (Chen et al., 2015), os pesquisadores demonstraram a capacidade de CRISPR em produzir macacos geneticamente modificados por meio da injeção de mRNA de Cas9 e de sgRNAs para os genes *NROB1*, *PPAR-γ* e *RAG1* em zigotos. Foram obtidos animais nascidos vivos com mutações nos genes *PPAR-γ* e *RAG1*, mas não em *NROB1*. Essas mutações estavam presentes também na linhagem germinativa destes animais, indicando uma provável capacidade de transmissão deste genótipo para sua progênie. Posteriormente, foram investigadas características dos fetos abortados que apresentavam mutações em *NROB1* (Kang et al., 2015). Interessantemente, um destes fetos do sexo masculino reproduziu características das doenças hipoplasia adrenal congênita (AHC) e hipogonadismo hipogonadotrófico (HH) em humanos, como defeitos no desenvolvimento do córtex da adrenal, arquitetura testicular anormal e, ao contrário do modelo murino para estas doenças, não apresentou alterações nas células de Sertoli. Estas pesquisas corroboram a utilidade de CRISPR para a produção de modelos mais fiéis para o estudo de doenças humanas utilizando primatas não humanos.

6.12 Conclusões

Neste capítulo, buscamos apresentar uma gama diversa de estratégias de utilização da ferramenta CRISPR em modelos animais úteis para a pesquisa científica e para a resolução de problemas de interesse biomédico e biotecnológico. No entanto, o pesquisador não está sob hipótese alguma limitado às espécies listadas. O chamado “zoológico CRISPR” (Reardon, 2016) faz referência digna ao crescente e já diverso número de espécies animais que têm se mostrado tratáveis à técnica, e ao fato de que novas descobertas e utilizações devem aparecer rapidamente. Portanto, o espetacular e ainda crescente impacto dessa tecnologia se justifica pela sua eficiência e versatilidade: em primeiro lugar ela tornou extremamente rápidos experimentos outrora longos e tediosos mesmo em animais-modelos amplamente utilizados; em segundo lugar, e talvez de modo ainda mais importante a longo

prazo, ela tornou possível a realização desses experimentos em modelos antes recalcitrantes à manipulação genética.

Desde o estabelecimento da genética e biologia molecular, barreiras técnicas representadas pela impossibilidade de se modificar o DNA de diversas espécies impediram o avanço do conhecimento específico em animais de maior interesse prático. A descoberta e ampla aplicabilidade da técnica de CRISPR significou um rompimento quase imediato dessas barreiras, tornando o futuro outrora promissor da biotecnologia em diversas espécies animais uma realidade instantânea. Tamanho poder e facilidade de uso identificados em tão pouco tempo são indicadores de que a comunidade científica ainda está longe de compreender todo o potencial desta ferramenta, e é certo dizer que, no momento, as limitações à sua utilização residem principalmente na disponibilidade de recursos, barreiras éticas, e na criatividade e conhecimento dos pesquisadores.

7. Células humanas

7.1 Introdução

Estudos *in vitro* representam estratégias importantes na investigação de funções gênicas e de mecanismos moleculares de doenças genéticas. Modelos celulares permitem uma avaliação controlada das consequências moleculares quando um gene é alterado ou quando fatores externos são modulados. Testes farmacêuticos de toxicidade e análise do efeito de drogas na fisiologia celular são exemplos da importância destes modelos na pesquisa científica, ainda que não representem toda a complexidade de um organismo, algo também essencial para uma investigação completa. Tanto nos campos da pesquisa básica quanto da aplicada, a edição genômica de células humanas utilizando a técnica de CRISPR representa uma ferramenta revolucionária, visto que a possibilidade de se induzir alterações específicas no genoma de um tipo celular de interesse, de forma relativamente simples e eficiente, abre caminho para descobertas grandiosas sobre como nossos genes interagem com o ambiente e até mesmo como partes diferentes de cada um deles são importantes para o correto funcionamento do todo. O quebra-cabeças do genoma humano está ainda mais perto de ser solucionado com o advento desta tecnologia, havendo diversas possibilidades de aplicações. Nesta seção, discutiremos algumas delas e também que tipos de perguntas podem ser respondidas em modelos celulares humanos editados por meio de CRISPR.

Diversas publicações utilizando a tecnologia CRISPR em linhagens celulares vêm surgindo desde que sua aplicação foi descrita pela primeira vez. A indução de *knock-ins* (He et al., 2016), *knockouts* (Aparicio-Prat et al., 2015; Chu et al., 2015; Schumann et al., 2015), substituições alélicas (Byrne et al., 2014; Mandai et al., 2015), grandes deleções (Bauer et al., 2015), além de alterações em células germinativas (Chapman et al., 2015; Wu et al., 2015), foram relatadas visando responder a diferentes perguntas biológicas. Protocolos detalhados descrevendo diferentes metodologias de introdução da maquinaria CRISPR em células-tronco e linhagens transformadas também já foram publicados, e dois deles (Hendricks et al., 2015; Ran et al., 2015) estão sendo testados em nosso laboratório, com o objetivo de se induzir mutações no gene *FBN1* que, quando alterado, é responsável pela ocorrência da Síndrome de Marfan. O objetivo da indução de modificações que levassem à perda de função deste gene ou à deleção de um éxon específico em uma mesma linhagem de células-tronco

pluripotentes humanas (hPSCs) é o estudo de correlações genótipo-fenótipo pouco estabelecidas para a síndrome.

7.2 Substituição alélica

Visando também estudar correlações entre genótipo e fenótipo, Mandai e colaboradores (2015) descrevem uma estratégia de substituição alélica interessante em termos de estudo do papel de polimorfismos em doenças genéticas humanas. Os autores analisaram o papel de um polimorfismo intrônico no gene *STK39* (serina-treonina quinase 39) relacionados à susceptibilidade à hipertensão em caucasianos, como demonstrado em estudos de associação ampla do genoma (GWAS). Os GWAS são bastante importantes na procura de mutações-alvo cuja correlação com uma determinada característica existe no nível populacional, mas não foi estudada em termos moleculares. Para induzir a alteração pretendida, os pesquisadores utilizaram um par de Cas9 *nickases human-codon optimized* contendo um sinal de localização nuclear (*nuclear localization signal*, NLS) codificadas pelo plasmídeo hCas9_D10A (Addgene #41816 – Mali et al., 2013) sob o promotor de citomegalovírus (CMV) com o terminador timidina quinase (TK). As *nickases* são nucleases que geram quebras de simples fita no DNA sendo, portanto, mais específicas. Além da nuclease, gRNAs complementares à região polimórfica de *STK39* a ser substituída foram escolhidos e clonados no vetor gRNA_Cloning Vector (Addgene #41824 – Mali et al., 2013) sob o promotor U6 da RNA Polimerase III com o sinal de terminação “TTTTT” (U6/T₆). Uma sequência de oligodesoxinucleotídeos simples fita (ssODN) doadora com braços de homologia complementares às extremidades da região polimórfica também foi transfectada em conjunto com os plasmídeos citados. Desta forma, os autores geraram a substituição de um único nucleotídeo (rs3754777A>G) em células HEK293T. As células efetivamente alteradas foram selecionadas considerando-se a presença de um sítio de restrição diferencial gerado pela alteração alélica. Para estudar que tipos de alterações moleculares o polimorfismo gerava, os autores analisaram a taxa de expressão do gene editado em linhagens homozigotas ou heterozigotas e notaram que a substituição alélica intrônica provocava aumento da transcrição deste gene. Também foi visto que a quantidade de proteína fosforilada estava aumentada nas células editadas e que isso se devia ao aumento da transcrição gênica e não a um aumento na taxa de fosforilação com relação à quantidade total de proteína. Outras proteínas da via de sinalização da qual STK39 participa também estavam aumentadas e, considerando-se que a via em questão tem papel na regulação da pressão arterial, as alterações encontradas podem explicar o aumento da susceptibilidade à hipertensão encontrado em pessoas portadoras da variante estudada.

Outro estudo que demonstrou sucesso na substituição alélica foi o de Byrne e colaboradores (2014). Os autores promoveram a substituição de 2,7 kb de um dos alelos do gene humano *THY1* (CD90) em células-tronco pluripotentes induzidas humanas (iPSCs) por seu homólogo murino (*Thy1*). Para isso, construíram um vetor de expressão da enzima Cas9 *human-codon optimized* contendo um NLS a partir da clonagem deste gene entre um promotor EF1α e a sequência de poliadenilação do hormônio de crescimento bovino (bGH poli-A) no plasmídeo pCDNA3, e sintetizaram diferentes gRNAs (U6/T₆) cujos alvos eram regiões próximas às extremidades de *THY1*. Também foi construído um vetor de substituição contendo a sequência de *Thy1* flanqueada por braços de homologia complementares às regiões externas aos sítios de corte no gene humano. Foram observadas substituições monoalélicas e bialélicas, em baixas frequências, como normalmente observado para alterações mediadas por reparo dependente de homologia (HDR). Outras alterações também foram observadas,

como deleções da região-alvo de *THY1* sem posterior inserção de *Thy1* em um ou dois alelos. Surpreendentemente, quando apenas um gRNA (em vez de dois) tendo como alvo *THY1* foi utilizado, a substituição aconteceu com maior frequência. O método de entrega dos vetores foi a nucleoporação, e a utilização de plasmídeos doadores linearizados diminuiu a eficiência do processo, visto que este método é mais eficaz para a entrega de vetores circulares. Os autores também encontraram uma correlação entre tamanho dos braços de homologia e eficiência de substituição, chegando a um tamanho ótimo de 2 kb, além de determinar que a geração de uma quebra de dupla fita, em vez de duas, é uma estratégia melhor para evitar inversões ou deleções, sendo mais eficiente para substituições.

7.3 Deleção de sequências gênicas

Outra interessante aplicação da técnica CRISPR foi demonstrada por Aparicio-Prat e colaboradores (2015), que descrevem uma estratégia para o nocauteamento de sequências gênicas, incluindo as de longos RNAs não codificadores (lncRNAs). O estudo dos papéis biológicos deste tipo de RNA é dificultado pelo fato de experimentos de *knockdown* por RNAi serem frequentemente ineficientes nesses casos. A estratégia utilizada para nocaute foi a deleção de promotores de dois proto-oncogenes de lncRNAs em linhagens celulares transformadas, sendo um deles o *MALAT1*. A deleção de promotores para eficiente silenciamento gênico é uma ferramenta interessante e apresenta vantagens em comparação à estratégia de deleção de genes inteiros, como o fato de o alvo ser menor e não variar tanto de tamanho como acontece com o corpo gênico total. A deleção de sequências menores pode garantir, com maior grau de certeza, que a perda de função do gene é a razão de uma eventual alteração de fenótipo, visto que a deleção de largas sequências pode levar a rearranjos não previstos. Outro destaque do protocolo é a clonagem de duas sequências de gRNA utilizando apenas um fragmento de DNA em um único plasmídeo lentiGuide-Puro (Addgene #52963 – Sanjana et al., 2014), já que o método frequentemente utilizado para a deleção de sequências é a indução de duas quebras de dupla fita, de forma que o fragmento de interesse seja excisado. O plasmídeo derivado, denominado pDECKO, continha dois diferentes gRNAs, um sob o promotor U6 e outro sob o promotor H1, além do terminador T₆. Para a efetiva edição, pDECKO foi transfectado nas linhagens celulares de interesse em conjunto com um vetor contendo o gene da Cas9 *human-codon optimized* com um NLS, o lentiCas9-Blast (Addgene #52962 – Sanjana et al., 2014), expresso sob o promotor EFS e com terminador bGH.

Quando o promotor de *MALAT1* foi deletado, observou-se redução quase total da expressão de seu lncRNA, o que demonstra a eficiência da estratégia. Além disso, também foi visto que os mutantes apresentavam redução das taxas de proliferação celular, o que é condizente com o papel oncogênico esperado para *MALAT1*.

Apesar das desvantagens descritas acima, a deleção de largas sequências genômicas pode ser essencial em alguns casos e trabalhos, como o de Bauer e colaboradores (2015) demonstram a factibilidade dessa estratégia utilizando a técnica de CRISPR. Os autores induziram a perda da função do gene *Pim1*, de tamanho maior que 5 kb, por meio de sua deleção completa e, para isso, desenharam dois gRNAs cujas sequências-alvo flanqueavam o gene de forma que um fragmento de tamanho total de 8 kb fosse excisado quando as quebras de dupla fita fossem geradas. Os gRNAs foram clonados no plasmídeo pX330 (Addgene #42230 – Cong et al., 2013) que já contém U6/T₆. Este vetor também contém o gene codificador da Cas9 *human-codon optimized* com NLS, sob o promotor CBh e com terminador bGH. Para análise do sucesso do nocauteamento, eles verificaram a perda da expressão do mRNA de

Pim1 por RT-PCR quantitativo e observaram ausência total de expressão nos clones com deleção bialélica. Apesar de esta mutação ter sido induzida em uma linhagem celular de camundongo (MEL), a estratégia é válida para células de mamíferos em geral.

7.4 Inativação gênica

O nocauteamento gênico também foi feito por Chu e colaboradores (2015), mas utilizando outra metodologia. O objetivo dos autores era investigar a função do gene *MUC18* em uma cultura primária de células epiteliais de vias aéreas, extremamente importantes na defesa do corpo contra partículas prejudiciais, como poluentes e patógenos. *MUC18* já foi descrito como potencial promotor de metástases quando superexpresso em células humanas de melanoma maligno. Sendo um gene codificador de uma proteína da superfamília das imunoglobulinas, os autores decidiram investigar seu papel inflamatório em culturas primárias de células epiteliais de via aérea e, para isso, utilizaram a estratégia de indução de apenas uma quebra de dupla fita imediatamente a 3' do códon de iniciação (primeiro ATG), de forma que o reparo por junção não homóloga de extremidades (NHEJ) resulte na formação de deleções e inserções (*indels*) e consequente mudança de quadro de leitura, nocauteando funcionalmente o gene. O método de entrega da maquinaria CRISPR foi o lentivírus contendo o vetor lentiCRISPRv1 (Shalem et al., 2014), que carrega os genes codificadores do gRNA (U6/T_e) e da Cas9 *human-codon optimized* contendo um NLS sob o promotor EFS e com elemento regulatório pós-transcricional WPRE. Sendo este um método integrativo, a expressão do gRNA, da Cas9 e também do gene de resistência à puomicina codificado pelo mesmo vetor torna-se constitutiva e estável. As quebras de dupla fita podem acarretar em diferentes alterações em diferentes células, visto que a geração de *indels* por reparo não homólogo é aleatória, e, por isso, a população resistente à puomicina selecionada é heterogênea quanto à alteração específica em *MUC18*, havendo células com efetiva perda de função gênica e células cujo *indel* gerado não acarretou em mudança do quadro de leitura. Análise da expressão do mRNA de *MUC18* nas linhagens editadas provenientes de dois diferentes doadores demonstrou redução de 93% e 84% quando comparados aos seus respectivos controles. Também não foi detectada a expressão da proteína MUC18 por *western blot*. O papel proinflamatório de MUC18 foi avaliado por meio da estimulação das linhagens editadas com agonistas de receptores *toll-like*, que mimetizam componentes virais e bacterianos. Foi observado que a produção da interleucina-8 (IL-8), uma das principais citocinas proinflamatórias, estava significativamente reduzida nas células nocaute para *MUC18*. Apesar de ser um exemplo de sucesso quanto à aplicação da técnica CRISPR na indução de alterações em células humanas primárias, a estratégia utilizada no trabalho apresenta algumas limitações, como apontado pelos próprios autores. Uma delas é a obtenção de uma população heterogênea de células editadas, em que proteínas funcionais ou com alterações podem ainda estar sendo expressas, acarretando em prejuízo para a análise de perda de função. Além disso, a expressão constitutiva da maquinaria de CRISPR gerada pela integração pode ocasionar mudanças na fisiologia celular que não necessariamente se relacionam ao nocaute específico de *MUC18*. Ainda assim, considerando-se a dificuldade de se trabalhar com culturas celulares primárias, o trabalho demonstra-se bastante importante no campo de edição genômica.

Schumman e colaboradores (2015) também descrevem o nocauteamento gênico em uma linhagem celular primária, mas usando uma metodologia diferente de entrega da maquinaria de edição genômica. Os autores utilizaram uma ribonucleoproteína (RNP) Cas9,

ou seja, uma Cas9 complexada a um gRNA transcrito *in vitro*, que tinha como alvo a região exônica do gene *CXCR4*. A transcrição do gRNA foi feita utilizando-se a T7 RNA polimerase e um molde de DNA contendo a sequência de interesse sob o promotor T7. A Cas9 foi expressa em bactérias *Escherichia coli* a partir do plasmídeo pMJ915 (Addgene #69090 – Lin et al., 2014), que contém o gene codificador da enzima com dois NLS sob o promotor e terminador T7. Após purificação, a nuclease foi complexada com o gRNA sintetizado. O gene *CXCR4* codifica um correceptor utilizado pelo vírus HIV para sua entrada em células T CD4⁺. Após eletroporação de células T primárias com a RNP Cas9, análise por citometria de fluxo revelou a existência de populações diferentes quanto à expressão de CXCR4, sendo a baixa expressão da proteína condizente de forma dose-dependente com a quantidade de ribonucleoproteína introduzida. Análise de *pool* de células das diferentes populações demonstrou a existência de mutações no gene *CXCR4* apenas no grupo com baixa expressão que recebeu a maquinaria de edição, enquanto o mesmo não foi observado para células controle. De forma interessante, foi visto que populações celulares eletroporadas que mantiveram alta expressão de CXCR4 não apresentavam mutações de terminação prematura da sequência codificadora, mas somente deleções *in frame* e mutações *missense*. Os autores também testaram a eficiência do método RNP Cas9 na substituição alélica mediada por HDR. Testando diferentes concentrações de uma sequência ssODN doadora que visava substituir parte do gene *CXCR4* de forma que a sequência-alvo do gRNA utilizado fosse deletada e substituída por um sítio de restrição de *HindIII*, eles observaram altas taxas de perda de expressão de CXCR4.

7.5 Integração de sequências exógenas (*knock-in*)

Ao contrário dos trabalhos descritos acima, He e colaboradores (2016) descrevem um método de geração de *knock-ins* gênicos independente do mecanismo HDR que, como mencionado, costuma ser pouco eficiente. Utilizando vetores doadores contendo uma sequência codificadora de GFP não flanqueada por braços de homologia, os autores induziram, com sucesso, a inserção do gene repórter no genoma celular. A ausência de braços de homologia no vetor doador impede que o mecanismo HDR seja ativado quando a quebra de dupla fita é gerada, sendo recrutada, portanto, apenas a maquinaria NHEJ. Para a indução da edição, foram utilizados vetores plasmidiais de expressão da Cas9 *human-codon optimized* convencional, hCas9 (Addgene #41815), ou *nickase*, hCas9_D10A (Addgene #41816), ambos de Mali et al. (2013). Como já mencionado para o segundo, o promotor utilizado para expressão da nuclease nestes vetores é o CMV e o terminador é o TK, além de a sequência gênica possuir um NLS. Também foram utilizados vetores de expressão de gRNAs específicos para uma região do gene *GAPDH* sintetizados de forma a conter U6/T_g, além de vetores contendo uma sequência eGFP flanqueada ou não por braços de homologia complementares à região-alvo. O objetivo era a indução de quebras de dupla fita no gene endógeno e a inserção da sequência eGFP na mesma região por HDR ou NHEJ, de forma que a célula editada passasse a ser GFP-positiva. Quando o vetor contendo braços de homologia foi utilizado, o mecanismo de reparo acionado para que a sequência eGFP fosse inserida deveria ser o HDR. Os autores observaram uma eficiência de menos de 1% de expressão de GFP em uma linhagem de células-tronco embrionárias humanas (hESCs) utilizando este método. A eficiência em células transformadas (LO2 e HEK293T), apesar de maior, foi de, no máximo, 5,97%. Quando uma Cas9 *nickase* foi utilizada, a eficiência de inserção do gene eGFP foi ainda menor. Entretanto, quando foi utilizado um vetor doador com uma sequência eGFP não flanqueada por braços homólogos a regiões do gene *GAPDH*, mas contendo também um ou dois gRNAs cujo alvo era o próprio plasmídeo

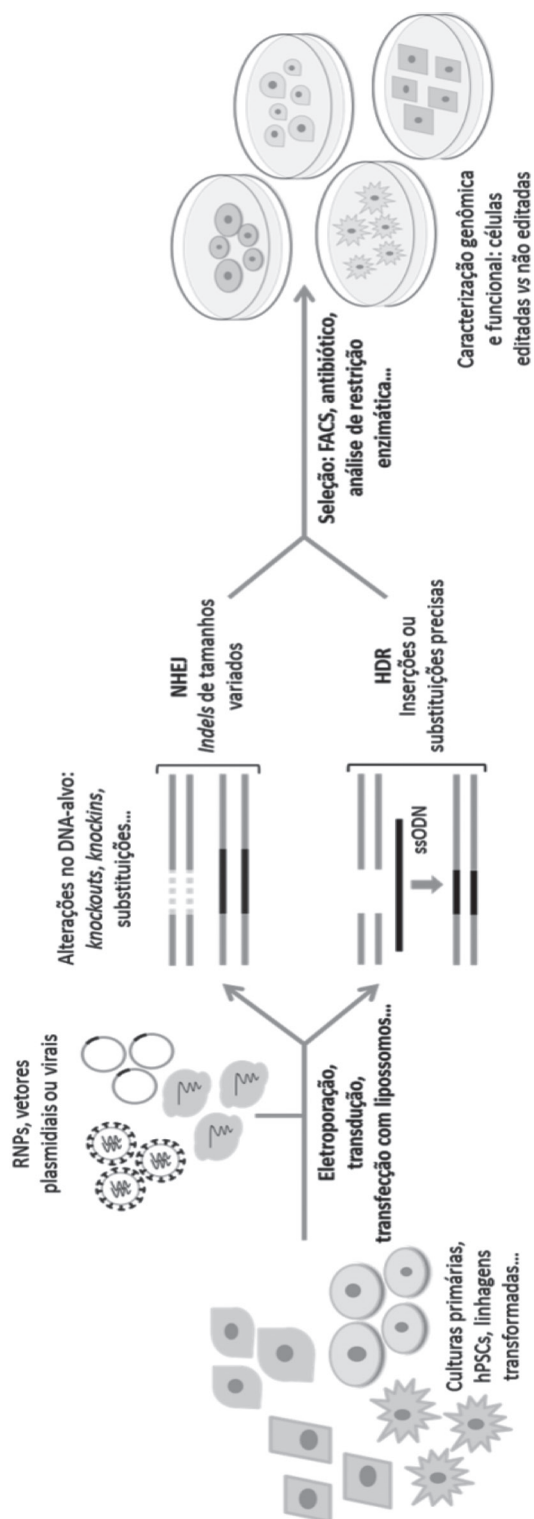


Figura 4. Diferentes estratégias para edição genômica em células humanas utilizando CRISPR. hPSCs: Células-tronco pluripotentes humanas; RNPs: ribonucleoproteínas; NHEJ: junção não homóloga de extremidades; *indels*: inserções ou deleções; ssODN: oligodesoxinucleotídeos simples fita; HDR: reparo mediado por homologia. FACS: separação de células ativada por fluorescência. Imagem: Borsoi, J e Pereira, LV (2016).

doador, foi observado um aumento bastante significativo da eficiência de *knock-in* (mais de 15% comparado a 5,97% em células LO2). Surpreendentemente, quando dois gRNAs cujos alvos eram cada uma das extremidades da sequência eGFP do plasmídeo doador foram utilizados, a eficiência foi menor que quando apenas um gRNA deste tipo foi inserido. Os autores testaram este sistema também para a inserção do GFP repórter em genes silenciados e obtiveram sucesso. Da mesma forma, comprovaram que a estratégia de *knock-in* mediado por NHEJ funciona em hESCs, apesar de em frequência muito mais baixa que em LO2.

7.6 Edição genética em espermatogônias

Dentre outros tipos celulares que já foram alvo de edição genômica utilizando a técnica CRISPR estão as espermatogônias de ratos. Chapman e colaboradores (2015) nocautearam dois genes, entre eles o *Epsti1*, visando uma prova de princípio de que mutações geradas neste tipo celular usando essa tecnologia seriam herdadas. O método utilizado para nocaute de *Epsti1* foi a eletroporação do vetor codificador da enzima Cas9 pX330 (Cong et al., 2013 – já mencionado) clonado com gRNAs específicos para o gene de interesse junto com um vetor portador do gene *Neo*, que confere resistência ao antibiótico geneticina (G418). As células submetidas ao processo foram, ou não, selecionadas e posteriormente transplantadas em túbulos seminíferos de ratos recipientes. Após pareamento com fêmeas selvagens para o gene, foram gerados animais mutantes heterozigotos tanto nos casos de espermatogônias selecionadas quanto com as não selecionadas. Porém, as proles de animais recipientes transplantados com células editadas que foram selecionadas quanto à resistência ao antibiótico geraram três vezes mais mutantes. Mutações de diferentes tipos foram observadas, sendo 56% de mudança de quadro de leitura e as restantes, deleções *in frame*. Dessa forma, foi comprovada a transmissão de mutações geradas em células germinativas utilizando CRISPR. Apesar de os experimentos descritos no trabalho terem sido feitos em células não humanas, demonstram a possibilidade de utilização desses modelos, junto com a tecnologia de edição genômica, para o estudo de genes relacionados à infertilidade em humanos que possuem homólogos em roedores, por exemplo.

As diferentes estratégias descritas são apenas algumas das muitas que já foram e ainda serão desenvolvidas usando técnicas de edição genômica. A figura 4 ilustra algumas delas. As possibilidades de perguntas biológicas que podem ser respondidas empregando esta tecnologia são inúmeras e colaboram cada vez mais para que avanços nos campos da Genética e da Biologia Molecular sejam alcançados.

Referências

- Acker J, Conesa C, Lefebvre O. Yeast RNA polymerase III transcription factors and effectors. *Biochim. Biophys. Acta*. 2013;1829:283-95.
- Aparicio-Prat E, Arnán C, Sala I, Bosch N, Guigó R, Johnson R. DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs. *BMC Genomics*. 2015;16(1):846.
- Asadollahi MA, Maury J, Schalk M, Clark A, Nielsen J. Enhancement of farnesyl diphosphate pool as direct precursor of sesquiterpenes through metabolic engineering of the mevalonate pathway in *Saccharomyces cerevisiae*. *Biotechnol Bioeng*. 2010;106:86-96.
- Associação Brasileira de Transplantes de Órgãos – ABTO. Anuário Estatísticas da ABTO. 2015. Available from: <http://www.abto.org.br/>
- Bakondi B, Lv W, Lu B, Jones MK, Tsai Y, Kim KJ, et al. In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa. *Mol Ther*. 2016;24(3):556-63.
- Bao Z, Xiao H, Liang J, Zhang L, Xiong X, Sun N, et al. Homology-Integrated CRISPR–Cas (HI-CRISPR) System for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth Biol*. 2015;4:585-94.

- Bassalo MC, Garst AD, Halweg-Edwards AL, Grau WC, Domaile DW, Mutalik VK, Arkin AP, Gill RT. Rapid and Efficient One-Step Metabolic Pathway Integration in *E. coli*. *ACS Synth. Biol.* 2016. <http://dx.doi.org/10.1021/acssynbio.5b00187>
- Bauer DE, Canver MC, Orkin SH. Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. *J Vis Exp.* 2015;(95). <http://dx.doi.org/10.3791/52118>.
- Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffin LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 2013;41:429-7437.
- Botstein D, Fink GR. Yeast: an experimental organism for 21st Century biology. *Genetics.* 2011;189:695-704.
- Brooks C, Nekrasov V, Lippman ZB, Van Eck J. Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiol.* 2014;166:1292-7.
- Byrne SM, Ortiz L, Mali P, Aach J, Church GM. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.* 2015;43(3):e21.
- Cha RS, Zarbl H, Keohavong P, Thilly WG. Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. *PCR Methods Appl.* 1992;2:14-20.
- Chapman KM, Medrano GA, Jaichander P, Chaudhary J, Waits AE, Nobrega MA, et al. Targeted germline modifications in rats using CRISPR/Cas9 and spermatogonial stem cells. *Cell Reports.* 2015;10(11):1828-35.
- Chen X, Rosbash M. mir-276a strengthens *Drosophila* circadian rhythms by regulating *timeless* expression. *Proc Natl Acad Sci USA.* 2016;113:E2965-72.
- Chen Y, Cui Y, Shen B, Niu Y, Zhao X, Wang L, et al. Germline acquisition of Cas9/RNA-mediated gene modifications in monkeys. *Cell Res.* 2015;25(2):262-5.
- Chu HW, Rios C, Huang C, Wesolowska-Andersen A, Burchard EG, O'Connor BP, et al. CRISPR-Cas9-mediated gene knockout in primary human airway epithelial cells reveals a proinflammatory role for MUC18. *Gene Ther.* 2015;22(10):822-9.
- Cobb RE, Wang Y, Zhao H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth Biol.* 2015;4:723-8. <http://dx.doi.org/10.1021/sb500351f>.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science.* 2013;339:819-23. <http://dx.doi.org/10.1126/science.1231143>.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819-23.
- Coradetti ST, Craig JP, Xiong Y, Shock T, Tian C, Glass NL. Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. *Proc Natl Acad Sci USA.* 2012;109:7397-402.
- Datta S, Costantino N, Court DL. A set of recombinering plasmids for gram-negative bacteria. *Gene.* 2006;379:109-15. <http://dx.doi.org/10.1016/j.gene.2006.04.018>.
- Di Donato V, De Santis F, Auer T, Testa N, Sánchez-Iranzo H, Mercader N, et al. 2C-Cas9: a versatile tool for clonal analysis of gene function. *Genome Res.* 2016. gr.196170.115.
- Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* 2013;41:4336-43.
- Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol.* 2014;32:1262-7. <http://dx.doi.org/10.1038/nbt.3026>.
- Doudna JA and Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014. Nov 28;346(6213):1258096
- Duina A, Miller ME, Keeney JB. Budding yeast for budding geneticists: a primer on the *Saccharomyces cerevisiae* model system. *Genetics.* 2014;197:33-48.
- Esvelt KM, Smidler AL, Catteruccia F, Church GM. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife.* 2014:e03401.
- Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, et al. Efficient CRISPR/Cas9-mediated targeted mutagenesis in *Populus* in the first generation. *Sci Rep.* 2015;5:12217.
- Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, et al. Gene targeting by the TAL effector PthXo2 reveals cryptic resistance gene for bacterial blight of rice. *Plant J.* 2015;82:632-43.
- Fausser F, Roth N, Pacher M, Ilg G, Sanchez-Fernandez R, Biesgen C, et al. *In planta* gene targeting. *Proc Natl Acad Sci USA.* 2012;109:7535-40.
- Fausser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* 2014;79:348-59.
- Galazka JM, Tian C, Beeson WT, Martinez B, Glass NL, Cate JHD. Cellodextrin transport in yeast for improved biofuel production. *Science.* 2010;330:84-6.
- Gantz VM, Bier E. The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations. *Science.* 2015;348:442-4.
- Gao J, Wang G, Ma S, Xie X, Wu X, Zhang X, et al. CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol Biol.* 2015;87:99-110.
- Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J Integr Plant Biol.* 2014;56:343-9.
- Garcia-Bloj B, Moses C, Blancafort P. The CRISPR road: from bench to bedside on an RNA-guided path. *Ann Transl Med.* 2015.
- Gems D, Johnstone IL, Clutterbuck AJ. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene.* 1991;98:61-7.
- Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 2009;6:343-5. <http://dx.doi.org/10.1038/nmeth.1318>.
- Gimble FS, Moure CM, Posey KL. Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system. *J Mol Biol.* 2003;334:993-1008.
- Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *MBio.* 2014;5:e00928-13. <http://dx.doi.org/10.1128/mBio.00928-13>.

- Harismendy O, Gendrel CG, Soularue P, Gidrol X, Sentenac A, Werner M, et al. Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J*. 2003;22:4738-47.
- He X, Tan C, Wang F, Wang Y, Zhou R, Cui D, et al. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res*. 2016. <http://dx.doi.org/10.1093/nar/gkw064>.
- Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 2014;11:122-3. <http://dx.doi.org/10.1038/nmeth.2812>.
- Hendriks WT, Jiang X, Daheron L, Cowan CA. TALEN and CRISPR/Cas9-Mediated Gene Editing in Human Pluripotent Stem Cells Using Lipid-Based Transfection. *Current Protocols in Stem Cell Biology*. 2015;34:5B.3.1-25.
- Heo YT, Quan X, Xu YN, Baek S, Choi H, Kim NH, et al. CRISPR/Cas9 nuclease-mediated gene knock-in in bovine-induced pluripotent cells. *Stem Cells Dev*. 2015;24(3):393-402.
- Horwitz AA, Walter JM, Schubert MG, Kung SH, Hawkins K, Platt DM, et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Syst*. 2015;1(1):88-96.
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31:827-32.
- Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M. Implementation of the CRISPR-Cas9 system in fission yeast. *Nat Commun*. 2014;5:5344.
- Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M. Implementation of the CRISPR-Cas9 system in fission yeast. *Nat Commun*. 2014;5:5344.
- Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA7. Parrott, Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol*. 2015;15:16.
- Jakočiūnas T, Bonde I, Herrgård M, Harrison SJ, Kristensen M, Pedersen LE, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab Eng*. 2015a;28:213-22.
- Jakočiūnas T, Rajkumar AS, Zhang J, Arsovska D, Rodriguez A, Jendresen CB, et al. CasEMBLR: Cas9-facilitated multiloci genomic integration of in vivo assembled DNA parts in *Saccharomyces cerevisiae*. *ACS Synth Biol*. 2015b;4:1226-34.
- Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, et al. EasyClone: method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res*. 2014;14:238-48.
- Jeong YH, Kim YJ, Kim EY, Kim SE, Kim J, Park MJ, et al. Knock-in fibroblasts and transgenic blastocysts for expression of human FGF2 in the bovine β -casein gene locus using CRISPR/Cas9 nuclease-mediated homologous recombination. *Zygote*. 2015;22:1-15.
- Jiang D, Zhu W, Wang Y, Sun C, Zhang KQ, Yang J. Molecular tools for functional genomics in filamentous fungi: Recent advances and new strategies. *Biotechnol Adv*. 2013;31:1562-74.
- Jiang W, Cox D, Zhang F, Bikard D, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013;1-9. <http://dx.doi.org/10.1038/nbt.2508>.
- Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res*. 2013.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816-21. <http://dx.doi.org/10.1126/science.1225829>.
- Kang Y, Zheng B, Shen B, Chen Y, Wang L, Wang J, et al. CRISPR/Cas9-mediated Dax1 knockout in the monkey recapitulates human AHC-HH. *Hum Mol Genet*. 2015;24(25):7255-64.
- Katayama T, Tanaka Y, Okabe T, Nakamura H, Fujii W, Kitamoto K, et al. Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus *Aspergillus oryzae*. *Biotechnol Lett*. 2016;38(4):637-42.
- Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*. 2014;24:1012-9.
- Kling J. First US approval for a transgenic animal drug. *Nat Biotechnol*. 2009;27:302-4.
- Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, et al. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol*. 2013;31:688-91.
- Li Z, Liu ZB, Xing A, Moon BP, Koellhoffer JP, Huang L, Ward RT, Clifton E, Falco SC, Cigan AM. Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant Physiol*. 2015. Oct;169(2):960-70.
- Liang Z, Zhang K, Chen K, Gao C. Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J Genet Genomics*. 2014;41:63-8.
- Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*. 2014;3:e04766.
- Liu R, Chen L, Jiang Y, Zhou Z, Zou G. Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system. *Cell Discov*. 2015;1:15007.
- Loison G, Vidal A, Findeli A, Roitsch C, Balloul JM, Lemoine Y. High level of expression of a protective antigen of schistosomes in *Saccharomyces cerevisiae*. *Yeast*. 1989;5:497-507.
- Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han Y-C, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature*. 2014;516:423-7.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823-6. <http://dx.doi.org/10.1126/science.1232033>.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823-6.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823-6.
- Mandai S, Mori T, Sohara E, Rai T, Uchida S. Generation of Hypertension-Associated *STK39* Polymorphism Knockin Cell Lines With the Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 System. *Hypertension*. 2015;66(6):1199-206.

- Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG, van den Broek M, et al. CRISPR/Cas9: a molecular swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 2015;15:1-15.
- Matsuura T, Baek M, Kwon J, Hong C. Efficient gene editing in *Neurospora crassa* with CRISPR technology. Fungal Biol. Biotechnol. 2015;2(4):1-7.
- Mumberg D, Müller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 1995;156:119-22.
- Nelles DA, Fang MY, Aigner S, Yeo GW. Applications of Cas9 as an RNA-programmed RNA-binding protein. BioEssays. 2015;37:732-9.
- Ni W, Qiao J, Hu S, Zhao X, Regouski M, Yang M, et al. Efficient gene knockout in goats using CRISPR/Cas9 system. PLoS One. 2014;9:e106718.
- Ninomiya Y, Suzuki K, Ishii C, Inoue H. Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci USA. 2004;101:12248-53.
- Nødvig CS, Nielsen JB, Kogle ME, Mortensen UHA. CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS One. 2015;10:e0133085.
- O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature. 2014;516:263-6.
- Oh J-H, van Pijkeren J-P. CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. Nucleic Acids Res. 2014;42:e131-131. <http://dx.doi.org/10.1093/nar/gku623>.
- Oldenburg KR, Vo KT, Michaelis S, Paddon C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. Nucleic Acids Res. 1997;25:451-2.
- Özaydin B, Burd H, Lee TS, Keasling JD. Carotenoid-based phenotypic screen of the yeast deletion collection reveals new genes with roles in isoprenoid production. Metab Eng. 2013;15:174-83.
- Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, Lu T, et al. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in *Caenorhabditis elegans*. Genetics. 2014;198:1347-56.
- Piatek A, Ali Z, Baazim H, Li L, Abulfaraj A, Al-Shareef S, et al. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. Plant Biotechnol J. 2015;13:578-89.
- Punt PJ, van Biezen N, Conesa A, Albers A, Mangnus J, van den Hondel C. Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol. 2002;20:200-6.
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152:1173-83.
- Quan J, Tian J. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. Nat Protoc. 2011;6:242-51. <http://dx.doi.org/10.1038/nprot.2010.181>.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308.
- Reardon S. Welcome to the CRISPR zoo. Nature. 2016;531:160-3.
- Roche CM, Loros JJ, McCluskey K, Louise Glass N. *Neurospora crassa*: Looking back and looking forward at a model microbe. Am J Bot. 2014;101:2022-35.
- Ronda C, Maury J, Jakočiūnas T, Baallal Jacobsen SA, Germann SM, Harrison SJ, et al. CrEdit: CRISPR mediated multi-loci gene integration in *Saccharomyces cerevisiae*. Microb Cell Fact. 2015;14(1):1-11.
- Ronda C, Pedersen LE, Sommer MO, Nielsen AT. CRMAGE: CRISPR Optimized MAGE Recombineering. Sci Rep. 2016;1-11. <http://dx.doi.org/10.1038/srep19452>.
- Ruiz-Díez BA. Review: Strategies for the transformation of filamentous fungi. J Appl Microbiol. 2002;92:189-95.
- Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. eLife. 2014:e03703.
- Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32:347-55. <http://dx.doi.org/10.1038/nbt.2842>.
- Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014;11(8):783-4.
- Schaeffer SM, Nakata PA. CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field. Plant Sci. 2015;240:130-42.
- Schiml S, Fauser F, Puchta H. The CRISPR/Cas system can be used as nuclease for *in planta* gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J. 2014;80:1139-50.
- Schiml S, Puchta H. Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. Plant Methods. 2016;12:8.
- Schumann K, Lin S, Boyer E, Simeonov DR, Subramaniam M, Gate RE, et al. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. Proc Natl Acad Sci USA. 2015;112(33):10437-42.
- Servick K. Gene-editing method revives hopes for transplanting pig organs into people. Science. 2015. Available from: <http://www.sciencemag.org/news/2015/10/gene-editing-method-revives-hopes-transplanting-pig-organs-people>
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343(6166):84-7.
- Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol. 2013;31:686-8.
- Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. Nat Protoc. 2009;4:206-23. <http://dx.doi.org/10.1038/nprot.2008.227>.
- Shi S, Liang Y, Zhang MM, Ang EL, Zhao H. A highly efficient single-step markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. Metab Eng. 2016;33:19-27.
- Storici F, Durham CL, Gordenin D, Resnick M, a 2003. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc Natl Acad Sci USA. 2003;100:14994-9.

- Stovicek V, Borodina I, Forster J. CRISPR-Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Metab Eng Commun.* 2015;2:13-22.
- Takahashi JS, Hong H-K, Ko CH, McDearmon EL. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet.* 2008;9:764-75.
- Tan W, Proudfoot C, Lillo SG, Whitelaw CB. Gene targeting, genome editing: from Dolly to editors. *Transgenic Res.* 2016. In press.
- Tong Y, Charusanti P, Zhang L, Weber T, Lee SY. CRISPR-Cas9 Based Engineering of Actinomycetal Genomes. *ACS Synth Biol.* 2015;4:1020-9. <http://dx.doi.org/10.1021/acssynbio.5b00038>.
- Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, Chulow JS, et al. Cytotoxic chDROMosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet.* 2013;9:e1003454. <http://dx.doi.org/10.1371/journal.pgen.1003454>.
- Waltz E. Gene-edited CRISPR mushroom escapes US regulation. *NATNEWS.* 2016;532(7599):293.
- Wang F, Shi Z, Cui Y, Guo X, Shi YB, Chen Y. Targeted gene disruption in *Xenopus laevis* using CRISPR/Cas9. *Cell Biosci.* 2015;5:15.
- Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature.* 2009;460:894-8. <http://dx.doi.org/10.1038/nature08187>.
- Wang S, Zhang S, Wang W, Xiong X, Meng F, Cui X. Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. *Plant Cell Rep.* 2015;34:1473-6.
- Wang X, Yu H, Lei A, Zhou J, Zeng W, Zhu H, et al. Generation of gene-modified goats targeting MSTN and FGF5 via zygote injection of CRISPR/Cas9 system. *Sci Rep.* 2015;5:13878.
- Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol.* 2014;32:947-51.
- Whelan W, Gocke E, Manney T. The CAN1 locus of *Saccharomyces cerevisiae*: fine-structure analysis and forward mutation rates. *Genetics.* 1979;91:35-51.
- Woo JW, Kim J, Kwon SI, Corvalán C, Cho SW, Kim H, et al. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol.* 2015;33(11):1162-4.
- Wu Y, Zhou H, Fan X, Zhang Y, Zhang M, Wang Y, et al. Correction of a genetic disease by CRISPR-Cas9-mediated gene editing in mouse spermatogonial stem cells. *Cell Res.* 2015;25(1):67-79.
- Xie K, Zhang J, Yang Y. Genome-wide prediction of highly specific guideRNA spacers for CRISPR-Cas9-mediated genome editing in model plants and major crops. *Mol Plant.* 2014;7:923-6.
- Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, et al. CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* 2014;29:327.
- Xu R, Li H, Qin R, Wang L, Li L, Wei P, et al. Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice. *Rice.* 2014;7(1):5.
- Yang L, Guell M, Niu D, George H, Leshia E, Grishin D, et al. Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science.* 2015;350:1101-4.
- Zhang C, Meng X, Wei X, Lu L. Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in *Aspergillus fumigatus*. *Fungal Genet Biol.* 2016;86:47-57.
- Zhang GC, Kong II, Kim H, Liu JJ, Cate JHD, Jin YS. Construction of a quadruple auxotrophic mutant of an industrial polyploid *Saccharomyces cerevisiae* strain by using RNA-guided Cas9 nuclease. *Appl Environ Microbiol.* 2014;80:7694-701.
- Zhou H, Liu B, Weeks DP, Spalding MH, Yang B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* 2014;42(17):10903-14.
- Zou Q, Wang X, Liu Y, Ouyang Z, Long H, Wei S, et al. Generation of gene-target dogs using CRISPR/Cas9 system. *J Mol Cell Biol.* 2015;7(6):580-3.

Otimizações da técnica

Capítulo 10

Tiago Alves Jorge de Souza^{1,a}, Rafael Sanguinetti Czepielewski^{2,3,b}, Daiana Renck^{2,b},
Cristina Bonorino^{2,b}, Jayme Augusto de Souza-Neto^{4,5,c} e Tiago Campos Pereira^{1,6,d}

¹*Programa de Pós-Graduação em Genética, FMRP USP, SP - Brasil*

²*Lab. de Imunologia Celular e Molecular, IPB, PUC RS, RS - Brasil*

³*Lab. de Descoberta de Drogas, IQ, PUC RS, RS - Brasil*

⁴*Instituto de Biotecnologia, UNESP - Botucatu, SP - Brasil*

⁵*Dept.º de Engenharia de Bioprocessos e Biotecnologia, FCA, UNESP, SP - Brasil*

⁶*Dept.º de Biologia, FFCLRP - USP, SP - Brasil*

^a*Autoria da seção: 2*

^b*Autoria da seção: 3*

^c*Autoria da seção: 4*

^d*Autoria das seções: 5 e 6*

Estruturação do capítulo

1. Introdução
2. Inibição da via NHEJ
3. *Multiplex* CRISPR
4. Reação Mutagênica em Cadeia (MCR)
5. Espécies transgênicas para Cas9
6. Conclusões do capítulo

1. Introdução

Veremos neste capítulo algumas interessantes otimizações de CRISPR que visam a: (i) aumentar a eficiência de recombinação homóloga, por meio da inibição da via de NHEJ, (ii) edição de múltiplos alvos simultaneamente (*multiplex* CRISPR), (iii) geração eficiente de células e animais com mutações homozigóticas (MCR) e (iv) sistemas baseados na administração apenas de gRNA (espécies transgênicas para Cas9).

2. Inibição da via NHEJ

Introdução

O reparo de DNA é essencial para a manutenção da viabilidade celular, pois ele resguarda a integridade genômica. Dentre as várias lesões que podem ocorrer no DNA, destacam-se as quebras de fita dupla (DSBs – *Double Strand Breaks*), por serem potencialmente as mais deletérias, pois caso não sejam reparadas, elas podem resultar em perda de material genético (Vartak e Raghavan, 2015). Em cada célula de mamíferos ocorrem ao redor de 10 DSBs por dia. Essas quebras podem ocorrer de forma endógena ou por fontes exógenas, como a radiação ionizante (Burma et al., 2006; Vartak e Raghavan, 2015). As principais vias de reparo de DSBs são a HDR e NHEJ (Lieber, 2010; Wyman e Kanaar, 2006). O reparo por NHEJ ocorre durante todo o ciclo celular, enquanto o reparo HDR está tipicamente ativo na fase de S e G2, quando os cromossomos estão duplicados (Deriano e Roth, 2013; Mali et al., 2013; Wang et al., 2013; Yang H, et al., 2013). No processo HDR, o reparo se dá por meio de uma fita molde homóloga à região da quebra do DNA, permitindo que o reparo seja preciso. Já no caso de NHEJ, o reparo envolve a adição ou deleção de algumas bases, resultando em alterações na sequência nucleotídica (Lieber, 2010).

A compreensão dessas duas vias de reparo tem sido essencial para aumentar a eficiência de ferramentas de edição genômica como a CRISPR, pois várias dessas abordagens induzem DSBs no DNA a fim de introduzir no genoma um determinado fragmento de interesse ou possibilitar a substituição alélica. A eficiência dessa inserção/substituição alélica vai depender de qual dos dois principais mecanismos de reparo (NHEJ/HDR) for induzido em resposta às quebras (Maruyama et al., 2015). Mesmo na presença de um DNA doador, NHEJ é mais frequente do que HDR, e a inativação da primeira passa a ser uma estratégia interessante quando se objetiva edições genéticas precisas (tipicamente mediadas pela HDR) (Basu et al., 2015; Chu et al., 2015; Maruyama et al., 2015; Robert et al., 2015).

Estratégias de inibição

O processo de reparo NHEJ possui várias etapas, nas quais existem moléculas que executam papéis fundamentais (vide cap. 2, figura 4). As proteínas do complexo Ku (Ku70 e Ku80), por exemplo, são as responsáveis por se ligar às extremidades do DNA fragmentado e protegê-las da degradação aleatória durante o reparo. Essas proteínas interagem com as DNA PKcs e com a nuclease Artemis formando o complexo DNA-PK/Artemis, o qual rege a atividade endonucleolítica durante o reparo. A família das polimerases X (Pol μ e Pol λ) também tem um papel essencial nessa via de reparo, pois essas enzimas são as responsáveis pela polimerização das cadeias após as DSBs. Por fim, o complexo DNA ligase IV-XRCC4-XLF age promovendo os

ajustes finais para a junção da dupla fitas (Davis e Chen, 2013; Lieber, 2010; Ochi et al., 2015; Srivastava e Raghavan, 2015; Vartak e Raghavan, 2015) (vide cap. 2, Apêndice).

Apesar de alguns estudos utilizarem proteínas adenovirais para inativar a proteína p53, e consequentemente o reparo por NHEJ (Chu et al., 2015; Robert et al., 2015), a ampla maioria das abordagens tem se valido da inativação das proteínas da via NHEJ, visando favorecer HDR (Basu et al., 2015; Chu et al., 2015; Maruyama et al., 2015; Robert et al., 2015; Vartak e Raghavan, 2015). As estratégias para esse fim são inúmeras e envolvem, por exemplo, o uso da técnica de Interferência por RNA (RNAi) ou determinados compostos químicos para a inibição do complexo Ku, do complexo DNA-PK/Artemis ou da DNA ligase IV (Beumer et al., 2013; Lin et al., 2014; Yu et al., 2015).

Maruyama e colaboradores (2015), por exemplo, utilizaram a molécula Scr7 para inibir NHEJ em células e em embriões de camundongos. Essa molécula se liga à DNA ligase IV, diminuindo sua afinidade por DSBs e consequentemente inibindo a sua função (Srivastava et al., 2012). Maruyama propôs que a coinjeção do cassete CRISPR-Cas9 e Scr7 em zigotos de camundongos permitiria que eles se desenvolvessem normalmente e apresentassem eficiência aumentada na inserção de fragmentos de interesse em seus genomas (Maruyama et al., 2015). Para testar essa hipótese, a eficiência de Scr7 foi primariamente mensurada em culturas celulares humanas (A549 e MelJuSo), expressando Cas9 e o sgRNA contra o gene *TSG101*. Essas células foram então transfectadas com um plasmídeo doador, e posteriormente a expressão de Cas9 foi induzida. A análise da inserção no sítio-alvo demonstrou que Scr7 foi capaz de inibir NHEJ e aumentar a eficiência do reparo HDR em até três vezes nas células A549 e dezenove vezes nas células MelJuSo. Assim, o próximo passo foi analisar a capacidade de Scr7 em aumentar a eficiência de inserções genômicas mediadas por CRISPR-Cas9 em zigotos de camundongos. Para isso, o cassete CRISPR-Cas9 e Scr7 foram injetados em zigotos, e esses foram implantados em fêmeas pseudogravídas. Após o nascimento da prole, a inserção em diferentes locos foi analisada e, assim como nas linhagens celulares (A549 e MelJuSo), foi observada a inibição de NHEJ e o aumento da eficiência de reparo por HDR (Tabela 1). Outros trabalhos têm corroborado esses dados, apontando para o potencial de Scr7 em aumentar a eficiência de edições genéticas mediadas por sistemas CRISPR-Cas9 (Chu et al., 2015; Robert et al., 2015; Singh et al., 2015; Vartak e Raghavan, 2015).

Entretanto, outras moléculas são conhecidamente eficientes no bloqueio da via NHEJ, como os inibidores químicos NU7441 e a KU-0060648, os quais se demonstraram capazes também de induzir HDR. Essa constatação se deu a partir dos experimentos de Robert e colaboradores (2015), os quais começaram suas análises silenciando por meio de RNAi componentes da via NHEJ em células humanas (HEK293/17). Para isso, essas células foram inicialmente transduzidas com o repórter TLR (*Traffic Light Reporter*) e transfectadas com siRNAs contra DNA-PKcs e DNA-ligase IV. Posteriormente, vetores de expressão Cas9/sgRNA foram introduzidos nas células. A análise por citometria de fluxo revelou um silenciamento de aproximadamente três vezes da DNA ligase IV e das DNA-PKcs concomitante ao aumento de HDR (Robert et al. 2015). A partir desses resultados, foi avaliado se inibição química da ação dos mecanismos de reparo HDR seria capaz de produzir resultados similares aos observados após o silenciamento por RNAi. Para esse propósito, as células TLR/293 foram inicialmente transfectadas com vetores expressando Cas9 e sgRNAs específicos contra os locos *Rosa26* e *TLR*, e 16h depois elas foram expostas a NU7441 e KU-0060648. Os resultados demonstraram que essas moléculas são eficientes na inibição de NHEJ, sendo capazes de aumentar em até duas vezes o reparo por HDR (Tabela 1).

Tabela 1. Resumo das diferentes abordagens (citadas nessa seção) utilizadas para inibir NHEJ e aumentar HDR em edições genômicas mediadas por CRISPR-Cas9.

Molécula ou estratégia utilizada	Loco	Efeito na NHEJ (% em relação aos níveis normais)	Efeito na HDR	Referência
Scr7	<i>TSG101</i>	Inibição	Aumento de 3x	Maruyama et al., 2015
Scr7	<i>TSG101</i>	Inibição	Aumento de 19x	Maruyama et al., 2015
Scr7	<i>Kell</i>	0%	Aumento de 2x	Maruyama et al., 2015
RNAi contra Ligase IV ou DNA PKcs	<i>Rosa26, TLR</i>	50%	Aumento de 2-3x	Robert et al., 2015
NU7441 ou KU-0060648	<i>Rosa26, TLR</i>	45%	Aumento de 2x	Robert et al., 2015
RNAi contra Ku70 ou Ku80	<i>Rosa26, TLR</i>	63%	Aumento de 2x	Robert et al., 2015
Scr7	<i>Rosa26, TLR</i>	70%	Aumento de 2x	Robert et al., 2015
Adenovírus (proteínas E1B55K e E4orf6)	<i>Rosa26, TLR</i>	12,50%	Aumento de 3,5x	Robert et al., 2015
RNAi contra Ligase IV	<i>Rosa26</i>	60%	Aumento de 3x	Chu et al., 2015
Scr7	<i>Rosa26</i>	40%	Aumento de 5x	Chu et al., 2015
RNAi contra KU70 e Ligase IV	<i>Rosa26</i>	50%	Aumento de 4-5x	Chu et al., 2015
Adenovírus (proteínas E1B55K e E4orf6)	<i>Rosa26</i>	18%	Aumento de 8x	Chu et al., 2015

A fim de mensurar o potencial da utilização de NU7441 e KU-0060648 na inibição de NHEJ e estímulo de HDR, Maruyama e colaboradores (2015) realizaram uma análise comparativa com outros métodos descritos na literatura como siRNAs contra Ku70 e Ku80, inibidores da DNA-ligase IV (Scr7) e expressão ectópica das proteínas E1B55K e E4orf6 (inibidores de p53) do adenovírus 5. A supressão de Ku70 e Ku80 por siRNAs na linhagem celular 293/TLR foi seguida da introdução de Cas9/sgRNAs e resultou na diminuição de até 1,6 vez a atividade NHEJ, enquanto foi capaz de aumentar duas vezes a atividade de HDR. A inibição da ligase IV pela Scr7 também foi capaz de estimular o aumento de HDR em aproximadamente duas vezes. Já a cotransfecção de vetores expressando as proteínas E1B55K e E4orf6 do adenovírus 5 demonstrou ser a abordagem mais eficiente, pois diminuiu em até oito vezes a atividade de NHEJ e aumentou mais de três vezes a atividade de HDR. A análise comparativa entre essas diferentes metodologias realizada por Robert e colaboradores se baseou, entre outros, em um estudo prévio conduzido por Chu e colaboradores (2015). Nesse trabalho foram comparadas as eficiências de três abordagens distintas em células humanas e de camundongos: (i) inibição por RNAi ou por Scr7 da DNA ligase IV, (ii) inibição

por RNAi de KU70 e KU80 (iii) coexpressão das proteínas E1B55K e E4orf6 do adenovírus 4. A inibição da KU70 e da DNA ligase IV foi capaz de aumentar em até cinco vezes a eficiência de HDR, enquanto a estratégia que utilizou a coexpressão das proteínas E1B55K e E4orf6 foi capaz de aumentar a eficiência de HDR em até oito vezes e reduzir drasticamente o reparo via NHEJ (Tabela 1).

Favorecimento de HDR em plantas

O controle dos níveis de reparo por NHEJ e HDR é igualmente essencial para a manipulação precisa dos genomas de plantas, havendo muitos esforços nesse sentido (Schiml e Puchta, 2016). Nesse contexto, um grande passo foi o desenvolvimento do sistema I-SceI-based *in planta* que permitiu taxas de sucesso na edição genética precisa maiores do que 1% em *A. thaliana* sem a necessidade de altas taxas de transformação. Para esse método, três T-DNAs são utilizados: uma sequência doadora, uma sequência-alvo I-SceI e um sistema de expressão I-SceI. Tanto a sequência doadora quanto a alvo possuem dois sítios de reconhecimento de I-SceI flanqueando suas sequências cassete. Já o sistema de expressão possui um promotor que quando induzido leva a expressão da endonuclease I-Sce. A ativação dessa enzima promove simultaneamente a indução de DSBs, a liberação do cassete doador e a ativação do cassete-alvo. A integração da sequência-alvo ocorre então mediante o reparo das DSBs via HDR (Fauser et al. 2012). Assim, por meio desse sistema a planta pode ter o seu genoma editado de forma precisa em qualquer uma das células ao longo do seu desenvolvimento (Puchta e Fauser, 2013; Schiml e Puchta, 2016). Interessantemente, com a utilização da nuclease Cas9 foi possível melhorar a eficiência dessa abordagem substituindo a necessidade da utilização de três sequências por apenas uma contendo a sequência doadora e o cassete de expressão Cas9/sgRNA (Schiml et al., 2013) (vide cap. 9 - GT *in planta*).

Conclusões

Em suma, uma das principais limitações da integração de transgenes/substituição alélica por meio de CRISPR se devia ao reparo de DSBs ser frequentemente executado por NHEJ em vez de HDR. Entretanto, diversas abordagens desenvolvidas nos últimos anos têm solucionado esse entrave, permitindo edições genéticas precisas, tanto em células animais quanto vegetais (Maruyama et al., 2015; Schiml e Puchta, 2016; Vartak e Raghavan, 2015).

3. Multiplex CRISPR

Introdução

Uma das grandes vantagens da técnica de CRISPR é poder desenhá-la para interagir com uma sequência específica de DNA (loco) através da enzima Cas9 (ou similar) e de moléculas de RNA (gRNAs). Entretanto, CRISPR pode ser projetada para reconhecer diversas sequências, em genes diferentes, simultaneamente, o que é denominado de “*multiplexing*” ou “*multiplex CRISPR*”. Nesse caso, em um único procedimento experimental diversas Cas9 atuam em conjunto com múltiplos gRNAs. Tais gRNAs podem ser desenhados para guiar Cas9 para atuar em diferentes genes, ou ainda em partes distintas de um mesmo gene, ampliando a possibilidade de inativação (*knockout*) (Ran et al., 2013), inserção (*knock-in*)

(Platt et al., 2014) e ativação gênica (Koner mann et al., 2014). Ao se adicionar gRNAs diferentes, podemos direcionar complexos Cas9/gRNA diferentes para regiões genômicas (locos) variadas. Se, por exemplo, desejamos intervir concomitantemente em dois genes diferentes (dois locos), basta adicionar os respectivos gRNAs para que a Cas9 atue nestes. Essa natureza programável e irrestrita que exemplifica o enorme potencial dessa técnica.

Entretanto, assim como ocorre com toda nova tecnologia, ainda existem restrições para o seu uso. Atualmente, o principal limitante é a metodologia que introduz o complexo codificador de Cas9/sgRNA no núcleo celular, ao alcance do DNA. E uma das variáveis fundamentais é a dimensão (em pares de bases) dos vetores, sejam eles virais ou plasmidiais. Por exemplo, cada gRNA deve ser expresso por um promotor de RNA polimerase III (Pol III) distinto. Por isso, apesar de as sequências de gRNA serem pequenas, com cerca de 100 nucleotídeos, a inserção de diversos promotores e o sinal de terminação (seis timinas – T₆) acarreta no aumento do tamanho total do vetor (figura 1 A-B). Na prática, o maior limitante do “*multiplexing*” acaba sendo a nossa competência de transfectar células sem afetar a sua integridade. Ou seja, depende da capacidade de inserir o material genético (vetor/plasmídeo) de grande volume, limitando assim o número de gRNAs a serem adicionados nesse contexto (figura 1 B-C). Um exemplo para dimensionar essa contenção é observar que a Cas9 mais utilizada no momento (proveniente de *Streptococcus pyogenes*, ou SpCas9) é codificada por um gene de aproximadamente 4,2 kb (Cong et al., 2013). Os mais eficientes vetores para transfecção celular são os vetores virais, como o AAV (*adeno-associated virus*), e esses têm capacidade limite de empacotamento de DNA, ou seja, o volume de DNA que pode ser adicionado ao sistema, de aproximadamente 4,4 kb, restando pouco espaço para inserção de múltiplos gRNAs (Swiech et al., 2015). Todavia, novos métodos de clonagem possibilitam obtenção de construtos otimizados e mais enxutos, possuidores de múltiplos promotores e gRNAs (figura 1 C-D). Além disso, são essenciais a utilização de promotores Pol III pequenos, como U6, e a utilização de sequências bicistrônicas virais (IRES ou 2A), as quais permitem a expressão simultânea de duas proteínas a partir de um único promotor (Cas9 e GFP, por exemplo). A seguir, descreveremos diferentes abordagens para construção de vetores que possibilitem a edição simultânea de dois ou mais genes.

Estratégias para construção de vetores para *multiplex*

Levando essas limitações em consideração, pode-se planejar o “*multiplexed CRISPR*” através de diferentes estratégias.

A maneira mais simplificada para atingir dois genes é usar um vetor contendo Cas9 + 2 gRNAs (figura 1B), através de dois promotores Pol III independentes (Maddalo et al., 2014). Nessa configuração, consegue-se alvejar dois genes, ou duas porções de um mesmo gene. Porém, nesse sistema, a colocação de gRNAs adicionais fica restrita devido à potencial redução da eficácia de transfecção e da potencialidade para ser empregado em vetores virais.

Uma estratégia alternativa é a utilização de dois vetores: um carregando a Cas9 e outro expressando múltiplos gRNAs, transfectando-os simultaneamente (figura 1C). Sendo desenhado com um **vetor 1**: contendo Cas9 + gene de seleção positiva (puromicina, por exemplo) e/ou proteína fluorescente (GFP, por exemplo); e **vetor 2**: contendo promotores Pol III e gRNAs + gene de seleção positiva (diferente do vetor 1), e/ou proteína fluorescente. Após a transfecção é necessário garantir a identificação e a seleção de células cotransfectadas, por isso os marcadores de seleção (genes de resistência) ou fluorescentes devem ser diferentes

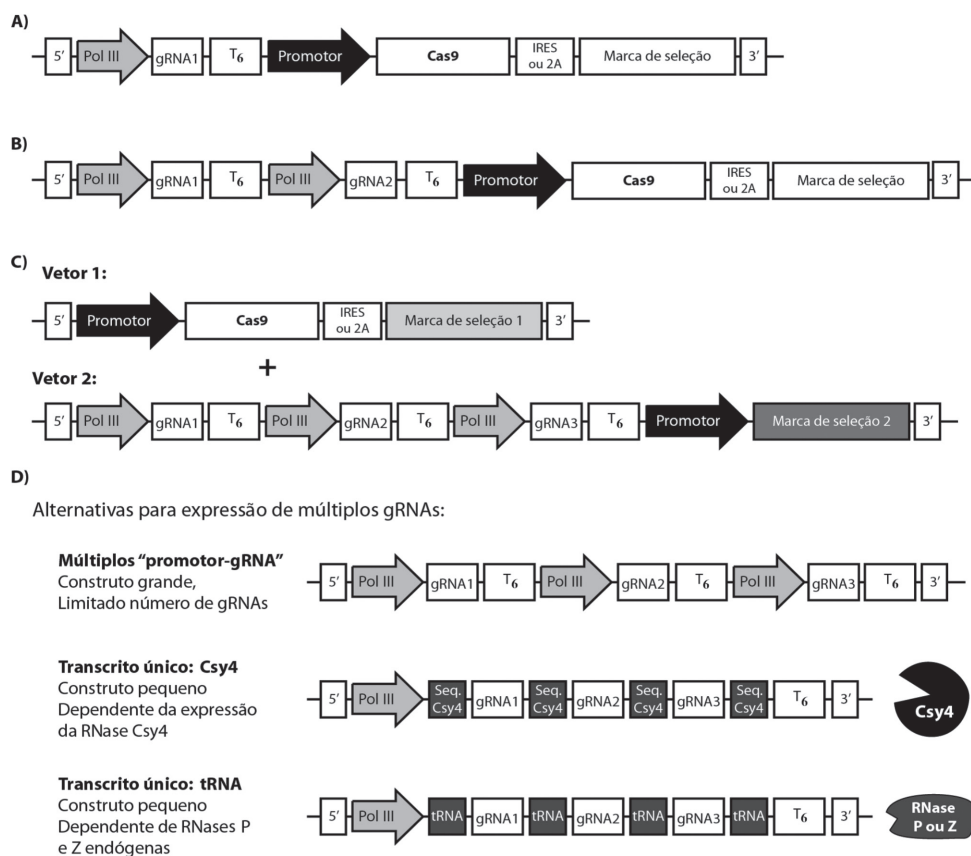


Figura 1. Alternativas de vetores para “multiplex CRISPR”. A-D) Estrutura esquemática dos vetores. Flechas simbolizam promotores de RNA Polimerase III (Pol III, cinza) e de RNA Polimerase II (preto). Retângulos representam sequências que codificam para gRNAs ou genes. T₆ simboliza sinal de terminação da transcrição. Sequências bicistrônicas IRES ou 2A permitem a expressão simultânea da Cas9 e marcador de seleção. “Marca de seleção” representa uma proteína fluorescente (GFP, por exemplo) ou um gene de resistência a antibióticos (puromicina, por exemplo). **A)** Construção de vetor para edição genética de um loco, com apenas um gRNA. **B)** Construção simples de vetor para edição genética de dois locos, com dois gRNAs. **C)** Estratégia com dois vetores: (1) expressando a Cas9 e (2) expressando múltiplos gRNAs. **D)** Alternativas para expressão de múltiplos gRNAs limitado pelo tamanho (superior) ou através de clivagem pela enzima Csy4 (meio), ou RNases endógenas P e Z (inferior). Imagem: Czieplewski, RS; Renck, D e Bonorino, C (2016).

para cada vetor. Essa configuração permite o emprego destes marcadores de confirmação e adicionam possibilidades experimentais para o seu modelo, como o advento das células editadas serem fluorescentes. Todavia, a adição de mais um vetor também irá reduzir a eficácia de edição gênica.

Outra abordagem é estabelecer um vetor único de maneira otimizada a expressar tanto Cas9 quanto múltiplos gRNAs (figura 1D). A construção dos vetores de todas as estratégias deve ser desenhada e realizada através de diferentes técnicas de biologia molecular, por exemplo: reações de clivagem e ligação enzimáticas (Ryan et al., 2014); ou através de protocolos otimizados de apenas uma etapa para ligação de múltiplos fragmentos, como “*Golden Gate Assembly*” (Kabadi et al., 2014; Sakuma et al., 2014; Xing et al., 2014), *USER* (Jakočinas et al., 2015), e *BioBrick* (Jiang et al., 2015; Shetty; Endy; Knight, 2008). A capacidade de “*multiplexing*” nesta

estratégia fica restrita ao número de gRNAs que podem ser clonados no vetor e apresentar atividade funcional.

Sendo assim, para permitir um maior número de gRNAs transcritos, foram usados métodos de clonagem especiais. Com a finalidade de obter de três a sete gRNAs, utiliza-se a estratégia de “*Golden Gate Assembly*” (Engler e Marillonnet, 2014). Esse método emprega uma enzima de restrição do tipo IIS que cliva o DNA após a sua sequência de reconhecimento (*BsaI*, por exemplo), criando extremidades livres. Essas extremidades podem ser desenhadas para conectarem-se ordenadamente, ligando múltiplos fragmentos que detenham essas sequências, culminando na inserção em outro vetor já estabelecido contendo Cas9 (figura 2 A-B). Para tanto, primeiramente, deve-se clonar os oligonucleotídeos de cada gRNA (gRNA1, gRNA2, gRNAⁿ, por exemplo) em vetores de expressão distintos através de uma enzima de restrição (*BpiI*, por exemplo) (figura 2A). Esses vetores devem conter sítios de restrição tipo IIS (reconhecidos pela *BsaI*) flanqueando a construção do promotor-gRNA (figura 2B). Após a clivagem, os diferentes fragmentos de promotor-gRNA podem ser ligados e inseridos em um vetor final com Cas9 (figura 2C). Similarmente, a construção de vetores virais pode ser efetuada através do método de “*Gateway Assembly*”. Este consiste em reações de recombinação desenhadas para inserir as diversas sequências de promotor-gRNA, através de sítios flaqueadores “*att*”, e utilizar eventos naturais de recombinação para conjugá-los em uma única construção (Albers et al., 2015; Buj et al., 2013).

Além disso, existem alternativas para produção de múltiplos gRNAs dependentes de um promotor único. Isso é obtido através da repetição de unidades de gRNA diferentes, em que o promotor transcreve tudo em uma única sequência de RNA. Para isso, os diferentes gRNAs devem estar flanqueados por sítios de clivagem, permitindo a liberação e processamento independente de cada gRNA. Isto reduz a necessidade de múltiplos promotores Pol III e, assim, diminui o vetor, permitindo também uma maior flexibilidade para utilização em vetores virais (figura 1D, meio). A obtenção desses gRNAs isolados provenientes de uma sequência única de gRNAs multiplexados pode ser realizada através de processos de clivagem pós-transcricional. Um utiliza a ribonuclease Csy4 (originária de *P. aeruginosa*) para clivar sítios específicos e liberar os diferentes gRNAs (Tsai et al., 2014). A desvantagem desse sistema é a necessidade da expressão, ou presença, de mais uma enzima nesse contexto. Entretanto, esse fato pode ser utilizado também como um fator de controle e regulação experimental, porque na ausência de Csy4 a nuclease Cas9 não será ativada (Nissim et al., 2014). A outra opção é utilizar sequências de RNA transportador (tRNA) flanqueando os gRNAs (figura 1D, inferior). O processamento endógeno do tRNA através das ribonucleases (RNases) P e Z retira as sequências transcritas de tRNA ao mesmo tempo que libera os gRNAs para atuarem (Mefferd et al., 2015; Xie; Minkenberg; Yang, 2015). Um terceiro sistema é a utilização de ribozimas intrínsecas, que são RNAs com capacidade catalítica de quebrar sequências de RNA (Scott; Horan; Martick, 2013). Ao flanquear gRNAs com sequências de ribozimas virais ou ribozimas *hammerhead* autocatalíticas são obtidos diversos gRNAs através de um mesmo transcrito (Nødvig et al., 2015; Yoshioka et al., 2015; Zhang; Matlashewski, 2015).

O uso de sistemas com promotor único para múltiplos gRNAs aparenta ser promissor, porém, estudos também demonstram que outro fator limitante para eficácia de edição do complexo é a quantidade e estabilidade dos gRNAs disponíveis (Hsu et al., 2013). Sendo assim, a expressão de gRNAs através de múltiplos promotores melhorará a sua produção, o que impacta a atividade final da Cas9, devendo ser priorizada quando se deseja editar poucas sequências de DNA.

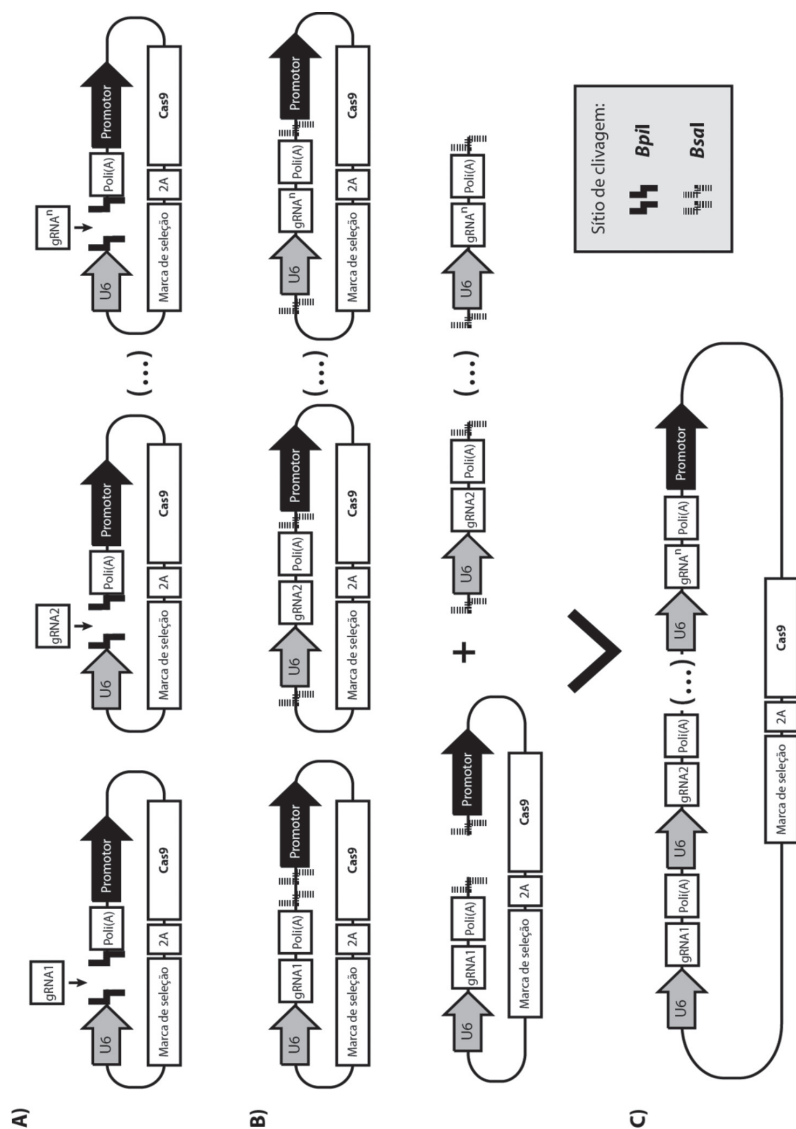


Figura 2. Estrutura esquemática do sistema de construção de vetores “Golden Gate Assembly”. A) Construção de múltiplos vetores CRISPR-Cas9 através da inserção de sequências de oligonucleotídeos (gRNA1, gRNA2 e gRNAⁿ) pela enzima de restrição *BpiI* (ou *BbsI*). B) Desenho esquemático do “Golden Gate Assembly”. O plasmídeo (superior à esquerda) apresenta um gRNA justaposto a uma sequência de reconhecimento da enzima *BsaI* antes do promotor da Cas9. Os plasmídeos (superiores, à direita) têm a sequência de *BsaI* flanqueando o promotor Pol III+gRNA. C) Após a clivagem com *BsaI* e ligação desses fragmentos obtemos um plasmídeo final contendo múltiplas sequências de promotor Pol III+gRNA distintos dentro do vetor que expressa Cas9 e marca de seleção. Imagem: Czepliewski, RS; Renck, D e Bonorino, C (2016). Adaptada de Sakuma et al., 2014.

Estratégias independentes de vetores

Considerando todas essas limitações, a microinjeção (no citoplasma de ovócitos, zigotos e células embrionárias) do mRNA da Cas9 e de gRNAs purificados vem sendo empregada na geração de animais geneticamente modificados (Niu et al., 2014; Wang et al., 2013, 2016). Esse processo não requer necessariamente a construção de vetores de transfecção, porém fica limitado aos eventos de microinjeção célula a célula e à pouca estabilidade do RNA. Todavia, permitem gerar edição genética *multiplex* em diversas espécies de seres vivos.

Uma síntese de alguns dos principais trabalhos recentes demonstrando a edição simultânea de vários genes em diversas espécies encontra-se na Tabela 2.

Conclusões

Em suma, o “*multiplex* CRISPR” tem mais virtudes do que ressalvas. Nos mais diferentes campos da pesquisa científica, essa ferramenta tende a ser cada vez mais empregada para alvejar múltiplos genes concomitantemente. Em vista disso, o desenvolvimento das técnicas de edição de múltiplos locos via CRISPR impactará todas as áreas da ciência, possibilitando um controle genético robusto para os que buscam responder às infinitas perguntas existentes.

4. Reação Mutagênica em Cadeia (MCR)

Introdução

Uma das mais promissoras variações da técnica CRISPR é a Reação Mutagênica em Cadeia (MCR, *Mutagenic Chain Reaction*). A MCR foi concebida por Gantz e Bier (2015) como um potente aprimoramento da mutagênese via CRISPR convencional, para a *conversão de mutações recessivas em heterozigose para a homozigose*, de forma eficiente e rápida, aplicável a células ou organismos (plantas ou animais).

Para compreender o efeito da MCR, considere os cenários a seguir. Imagine que se queira utilizar CRISPR para produzir uma linhagem celular para o estudo de determinada doença genética recessiva. Após se administrar Cas9/gRNA na cultura celular, a edição genética terá início. Algumas células terão ambos alelos editados (denominadas bialélicas), outras células terão apenas um alelo editado (denominadas monoalélicas) e algumas não serão editadas (selvagens). Note que as alterações celulares associadas à doença genética só emergirão em células bialélicas, que muito provavelmente apresentarão mutações diferentes em cada um dos dois alelos (*e.g.*, um alelo com inserção de 1 nt e o outro alelo com deleção de 4 nt). Ou seja, muito frequentemente as células bialélicas serão heterozigóticas. Um simples processo de seleção permitirá isolar estas células bialélicas, que provavelmente manifestarão o fenótipo associado à condição genética. Contudo, interações entre mutações distintas podem ser um complicador adicional, gerando efeitos inesperados.

Um problema semelhante, porém ainda mais complexo, ocorre na geração de animais-modelos para estas mesmas doenças genéticas recessivas. A edição via CRISPR em células embrionárias resultará em linhagens celulares monoalélicas, bialélicas ou selvagens. Note que o sucesso deste procedimento dependerá do fato de a linhagem germinativa se derivar de células bialélicas (resultando em 100% dos gametas editados) ou monoalélicas (50% dos gametas editados). Somente após cruzamentos controlados entre a prole F_1 (heterozigota) é que

Tabela 2. Síntese com alguns dos principais trabalhos de edição simultânea de vários genes em diversas espécies.

Espécie	Células/ Órgão	Nº locos editados	Tipo de edição	Forma de administração	Vetor	gRNA por vetor	Modelo	Ano	Referência
Bactérias									
Bactéria (<i>Escherichia coli</i>)	BL21 Star, BLΔsucC, MG1655	5	KO	Transformação	Único	5	Regulação do metabolismo	2015	Cress et al., 2015
Fungos									
Levedura (<i>Saccharomyces cerevisiae</i>)	CEN.PK2- 1C	5	KO	Transformação por acetato de lítio	Único	5	Regulação do metabolismo	2015	Jakočičinas et al., 2015
Levedura (<i>Saccharomyces cerevisiae</i>) e <i>Kluyveromyces lactis</i>	CEN.PK2 e K Lactis (ATCC 8585)	3	KO	Transformação	Único	3	Regulação do metabolismo	2015	Horwitz et al., 2015
Plantas									
Arabidopse (<i>Arabidopsis thaliana</i>) e milho (<i>Zea mays</i>)	Protoplasto	2	KO	<i>Agrobacterium</i> e PEG	Único	2	Geração de planta KO	2014	Xing et al., 2014
Arabidopse (<i>Arabidopsis thaliana</i>), Arroz (<i>Oryza sativa</i> <i>J. ssp</i>), Tabaco (<i>Nicotiana benthamiana</i>)	Protoplasto	3	KO	<i>Agrobacterium</i>	Único	3	Geração de planta KO	2015	Lowder et al., 2015
Arroz (<i>Oryza sativa</i> <i>L. ssp</i>)	Protoplasto	4	KO	<i>Agrobacterium</i>	Único	4	Geração de planta KO	2015	Xie et al., 2015
Soja (<i>Glycine max</i> L.)	Protoplasto	3	KO	<i>A. rhizogenes</i> e PEG	Múltiplos	1	Geração de planta KO	2015	Sun et al., 2015
Invertebrados									
Toxoplasma (<i>Toxoplasma gondii</i>)	Parasita RH88	2	KO	Eletroporação	Único	2	Geração de parasita KO	2016	Long et al., 2016
Nematódeo (<i>Caenorhabditis elegans</i>)	Verme adulto	3	KO	Microinjeção	Único	1	Geração de animal KO	2014	Kim et al., 2014
	Verme adulto	2	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2014	Chen et al., 2014

KO=knockout; KI=knock-in; CRISPR= ativação gênica.

* = Promega

**= Life Technologies

*** = Nepa Gene

Tabela 2. Continuação.

Espécie	Células/ Órgão	Nº locos editados	Tipo de edição	Forma de administração	Vetor	gRNA por vetor	Modelo	Ano	Referência
Invertebrados	Mosca (<i>Drosophila melanogaster</i>)	Embrões	4	KO	Microinjeção	Único	4	Geração de animal KO	2014 Port et al., 2014
	Mosquito (<i>Aedes aegypti</i>)	Embrões	3	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2015 Kistler et al., 2015
Peixes	Peixe-zebra (<i>Danio rerio</i>)	Embrões	6	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2015 Shah et al., 2015
		Embrões	3	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2015 Varshney et al., 2015
		Embrões	5	KO	Microinjeção	Único	5	Geração de animal KO	2015 Yin et al., 2015
Aves	Galinha (<i>Gallus gallus domesticus</i>)	PGC (células germinativas primordiais)	2	KO	Lipofectamine 2000**	Múltiplos	1	Geração de animal KO	2016 Oishi et al., 2016
		ESC, embrões	6	KO	FuGENE® HD*	Único	gRNAs isolados	Geração de animal KO	2013 Wang et al., 2013
Roedores	Camundongo (<i>Mus musculus</i>)	Células de medula óssea	5	KO	Lentiviral (pLKO-TRC005)	Múltiplos	1	Doença mieloide maligna	2014 Heckl et al., 2014
		Neurônios, células imunes e endoteliais	3	KI	AAV, lentiviral e nanopartícula	Único	3	Modelo genético tumoral	2014 Platt et al., 2014

KO=knockout; KI=knock-in; CRISPRa: ativação gênica.
* = Promega
**= Life Technologies
***= Nepa Gene

Tabela 2. Continuação.

Espécie	Células/ Órgão	Nº locos editados	Tipo de edição	Forma de administração	Vetor	gRNA por vetor	Modelo	Ano	Referência
Roedores	Fígado	10	KO	HTVI e Lipofectamine 2000**	Múltiplos	1	Modelo genético tumoral	2015	Weber et al., 2015
	Pâncreas	15	KO	Eletroporação <i>in</i> <i>vivo</i> NEPA21***	Múltiplos	1	Modelo genético tumoral	2016	Maresch et al., 2016
	Pulmão	2	KO	Adenoviral (Viraquest)	Único	2	Modelo genético tumoral	2014	Maddalo et al., 2014
	MEF	3	KO	Lentiviral	Único	3	Modelo genético tumoral	2015	Albers et al., 2015
	CHO	3	KO	FreeStyleTM (Thermo Fisher)	Único	3	Inativação de múltiplos genes	2015	Grav et al., 2015
Rato (<i>Rattus norvegicus</i>)	Embrões	4	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2014	Ma et al., 2014
Coelho (<i>Oryctolagus cuniculus</i>)	Embrões	5	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2014	Yan et al., 2014
Suínos	LDC, ovócitos, indivíduo	3	KO	Neon@**	Único	3	Geração de animal KO	2015	Li et al., 2015
	Zigoto, indivíduo	3	KO	Microinjeção	Múltiplos	gRNAs isolados	Geração de animal KO	2016	Wang et al., 2016

KO=knockout; KI=knock-in; CRISPRa: ativação gênica.

* = Promega

**= Life Technologies

***= Nepa Gene

Tabela 2. Continuação.

Espécie	Células/ Órgão	Nº locos editados	Tipo de edição	Forma de administração	Vetor	gRNA por vetor	Modelo	Ano	Referência
Suínos	Porco (<i>Sus scrofa domestica</i>)	62	KO	Lentiviral e Lipofectamine 2000**	Único	2	Remoção de vírus endógenos	2015	Yang et al., 2015
	Cabra (<i>Capra aegagrus</i>)	2	KO	Microinjeção	Múltiplos	gRNAs isolados	Geração de animal KO	2015	Wang et al., 2015
Caprinos	Macaco-cinomolgo (<i>Macaca fascicularis</i>)	3	KO	Microinjeção	Único	gRNAs isolados	Geração de animal KO	2013	Niu et al., 2013
	HEK 293FT	2	KO	Lipofectamine 2000**	Único	2	Inativação de múltiplos genes	2013	Cong et al., 2013
Primatas	HEK 293FT	4	Regulação genética	FuGENE® HD*	Único	4	Regulação genética	2014	Nissim et al., 2014
	OVCAR8-ADR	2	KO	Lentiviral (pLKO.1)	Múltiplos	2	Função combinatória de genes	2016	Wong et al., 2016
	HEK 293FT	4	KO e CRISPRa	Lentiviral e Lipofectamine 2000**	Único	4	Ativação de múltiplos genes	2014	Kabadi et al., 2014
	HEK 293FT	7	KO	Lipofectamina LTX**	Único	7	Inativação de múltiplos genes	2014	Sakuma et al., 2014
	HeLa e HEK293T	3	CRISPRa	FuGENE® HD * e Microinjeção	Múltiplos	1	Ativação de múltiplos genes	2013	Cheng et al., 2013

KO=knock-out; KI=knock-in; CRISPRa: ativação gênica.
* = Promega
**= Life Technologies
***= Nepa Gene

se obterão indivíduos homozigotos na F_2 . Note que, apesar de possível, é muito improvável que o animal F_0 (parental, derivado do embrião submetido à modificação via CRISPR) apresente o fenótipo da doença.

Estes dois cenários demonstram que a geração de células/indivíduos portando uma dada mutação recessiva *em heterozigose* é problemático, pois o alelo selvagem camufla o efeito do alelo mutante (em genótipos ^{mutante/selvagem}) ou pode ser um complicador (em genótipos ^{alelo mutante 1/alelo mutante 2}). Tal fato torna extremamente laboriosa a produção de linhagens mutantes homozigotas, pois exige seleção, múltiplos cruzamentos para a segregação mendeliana e consequente purificação da linhagem contendo o genótipo desejado.

A MCR é uma estratégia baseada em CRISPR que visa à obtenção de células/organismos mutantes homozigotos logo após o experimento de edição genética. Talvez o mais notável seja o fato de a MCR elevar consideravelmente as chances do próprio animal (ou planta) F_0 apresentar o fenótipo associado à condição genética recessiva.

O funcionamento da MCR pode ser resumido em três passos básicos (figura 3): no primeiro, ocorre a clivagem de um dado loco genômico através da ação da Cas9 direcionada

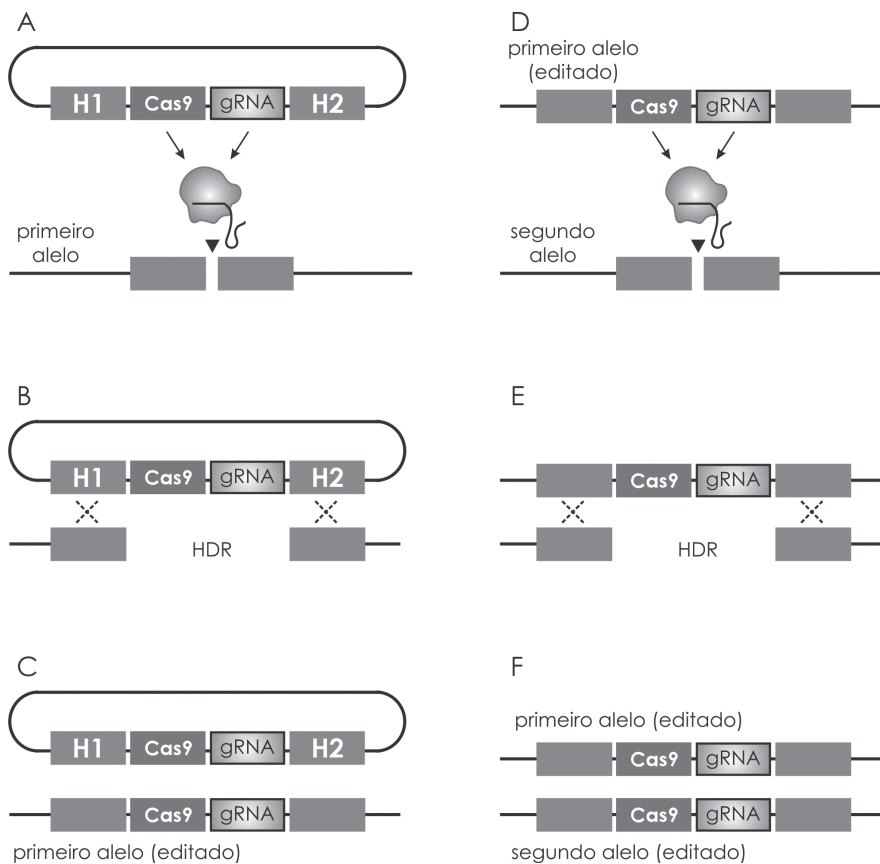


Figura 3. Princípio da Reação Mutagênica em Cadeia (MCR). A “construção MCR” (plasmídeo em A) permite a edição de um alelo de um loco específico do genoma. Este alelo endógeno, uma vez editado, promove a edição do segundo alelo endógeno, resultando em uma mutação homozigótica. HDR: via de reparo dependente de homologia. Baseado em Gantz e Bier, 2015. Imagem: Pereira, TC (2016).

por um gRNA específico, ambos expressos por um plasmídeo previamente incorporado pela célula; no segundo, a sequência de DNA exógena que codifica esses elementos é integrada à região genômica clivada através do sistema celular de reparo direcionado por homologia (HDR) (Jasin e Rothstein, 2003; Aylon e Kupiec, 2004); no terceiro, esse loco modificado, agora com *atividade autocatalítica*, é capaz de induzir o processo de clivagem e HDR no loco correspondente de seu cromossomo homólogo, transferindo assim a mutação para o mesmo. Para que isso seja possível, o plasmídeo utilizado inicialmente deve conter necessariamente a sequência de três elementos: i) o gRNA que será expresso e tem como alvo o loco genômico de interesse; ii) o gene que codifica a Cas9 sob o controle de um promotor que propicie sua expressão em ambas células somáticas e germinativas; e iii) fragmentos de DNA flanqueando os genes que codificam a Cas9 e o gRNA, e que necessariamente sejam homólogos a ambas extremidades adjacentes ao ponto de clivagem no genoma (Gantz e Bier, 2015).

Experimentos pioneiros envolvendo MCR realizados por Gantz e Bier (2015) utilizaram esse sistema expressando um gRNA específico para induzir mutações no gene *yellow* (*y*) do organismo modelo *Drosophila melanogaster*. Como o gene *y* está ligado ao cromossomo sexual X, indivíduos machos precisam de apenas uma cópia do cromossomo X com mutação nesse loco para exibirem o típico fenótipo de corpo com pigmentação amarelada. Por outro lado, para que o mesmo fenótipo seja visualizado em moscas fêmeas, a mutação deve estar presente em homozigose, uma vez que a mutação é de caráter recessivo.

O cruzamento entre moscas da geração F_0 , cujos embriões foram injetados com a construção “y-MCR”, e indivíduos do tipo selvagem (*y*+) foi capaz de gerar descendentes fêmeas com corpo amarelo (*y*-), o que contraria a herança mendeliana, segundo a qual todas as fêmeas deveriam apresentar ao menos um alelo *y*+ (oriundo da linhagem parental selvagem) e, portanto, corpo com coloração normal. A capacidade de conversão alélica mediada por MCR foi ainda mais notável quando uma fêmea da primeira geração de moscas mutantes foi cruzada com um macho selvagem (*y*+), resultando em ~97% dos descendentes apresentaram corpo com pigmentação amarelada tal qual o fenótipo *y*-.

MCR-Gene drives e o controle de doenças transmitidas por vetores

A técnica de MCR funciona de forma eficiente não apenas em organismos-modelos, tal qual *D. melanogaster*, mas também em insetos vetores de doenças como o mosquito transmissor da malária *Anopheles stephensi*. Nesse caso, Gantz et al. (2015) utilizaram MCR como um *sistema que propaga e fixa alelos/genes ao longo de poucas gerações*, conhecido como *gene drive system* (revisado por Sinkins e Gould, 2006). Neste caso, além dos componentes da MCR citados acima, a construção de DNA transformante possuía também um cassete de DNA transgênico com as sequências promotoras e codificadoras necessárias para a expressão de moléculas efetoras capazes de bloquear o desenvolvimento do parasita causador da malária no mosquito (Isaacs et al., 2011; Isaacs et al., 2012; Li et al., 2005). A injeção dessa construção em embriões de mosquitos resulta em uma taxa de conversão de heterozigotos em homozigotos em torno de 100%, ressaltando a alta eficiência da técnica. Além disso, a expressão dos transgenes efetores continua sendo estimulada de forma robusta e específica em determinados órgãos do inseto, demonstrando que a inserção mediada por MCR não altera o seu funcionamento (Gantz et al., 2015).

Estes resultados abrem novas e promissoras possibilidades para o controle não apenas da malária, mas de diversas outras doenças transmitidas por vetores, tais quais dengue e zika. *Gene drives* mediados por MCR poderiam ser utilizados, por exemplo, para a fixação,

em populações selvagens, de transgenes conferindo caráter de refratoriedade a um dado patógeno através da liberação de mosquitos geneticamente modificados (GMs) por MCR na natureza. Estudos de simulação e modelagem matemática atualmente se propõem a investigar e estimar a eficácia desta abordagem inovadora (Unckless et al., 2015).

Terapias gênicas

A edição de genomas mediada por CRISPR-Cas9 tem sido apontada como uma tecnologia muito promissora para o tratamento de doenças virais e genéticas através terapias gênicas (Meaders e Gersbach, 2016; Stone et al., 2016;). Nesse sentido, a MCR é apresentada por Gantz e Bier (2015) como uma ferramenta ainda mais eficiente, capaz de modificar o genoma não apenas de células germinativas mas também de células somáticas. Neste caso, a MCR poderia ser utilizada com um sistema de entrega mediado por vírus, o que em teoria possibilitaria o tratamento de quaisquer doenças que causem alterações na sequência genômica. Por exemplo, elementos MCR poderiam ser integrados ao gene que codifica a proteína Integrase de retrovírus que se inserem no genoma do hospedeiro, tal qual o HIV, para combater os reservatórios latentes do vírus e prevenir sua propagação para células T do sistema imune. Outra possibilidade seria a produção de vírus não patogênicos portando elementos MCR capazes de se integrar e propagar especificamente em regiões do genoma específicas de células cancerosas. O elemento vírus-MCR poderia ainda conter uma sequência que codificasse uma toxina induzida por drogas, o que causaria a eliminação programada e controlada de células tumorais apenas.

5. Espécies transgênicas para Cas9

Introdução

O aperfeiçoamento de protocolos é uma meta constante em qualquer área, buscando abordagens mais simples, eficientes, rápidas ou menos caras. Uma das primeiras e mais importantes otimizações em CRISPR foi a redução dos elementos necessários para edição genética, passando de três (Cas9, crRNA e tracrRNA) para apenas dois (Cas9 e sgRNA). Contudo, a manipulação de dois elementos ainda pode ser problemática em algumas situações. Por exemplo, a entrega de Cas9/gRNA *in vivo* (camundongos) é frequentemente baseada em vetores virais, que apresentam uma limitação física quanto à capacidade de DNA que pode ser empacotado dentro de suas estruturas. Considerando que a SpCas9 possui ~4,1 kpb, além dos promotores, terminadores e o sgRNA, não resta muito espaço para outros elementos (*i.e.*, GFP ou proteínas repressoras, para *DNA labeling* ou *regulação da expressão gênica*, respectivamente).

Se considerarmos que a Cas9 é um elemento *constante* em todos os experimentos de edição genética – já que variamos apenas o gRNA a cada procedimento –, por que não desenvolver uma estratégia em que a *Cas9 faça parte integral do sistema*, sendo necessário administrar apenas o gRNA? Em outras palavras: por que não criar células/organismos que expressem Cas9 de forma heteróloga, limitando-nos a introduzir apenas o gRNA a cada experimento? Pensando nisso foram criadas linhagens transgênicas para a nuclease Cas9.

Fungo Cas9

A construção de uma linhagem transgênica do fungo patogênico *Aspergillus fumigatus* expressando Cas9 (**AfCas9**) foi apresentada por Fuller e colaboradores em 2015. Neste trabalho, os autores introduziram um cassete de expressão de Cas9 (otimizada para códons humanos) sob controle do promotor *tef1* e terminador *cyc1* por transformação química de protoplastos (Fuller et al., 2015). De acordo com os pesquisadores, a expressão constitutiva de Cas9 nesta linhagem não alterou a taxa de crescimento, desenvolvimento, resistência a estresse oxidativo, homeostase da parede celular, sensibilidade a antifúngicos ou virulência no modelo murinho. Estes resultados demonstram que **AfCas9** pode ser usada como modelo dentro de uma plataforma de *screening* genético, em que a introdução de cassetes de expressão de diferentes gRNAs, via transformação química de protoplastos, permitirá uma análise funcional em larga escala.

Drosófila Cas9

Possivelmente, o primeiro animal transgênico para Cas9 foi a mosca-das-frutas, *Drosophila melanogaster*. Em novembro de 2013, uma dupla de japoneses publicou seus dados referentes a uma linhagem de drosófila com Cas9 (**DmCas9**) sob controle do promotor *nanos*, permitindo assim a expressão restrita da nuclease na linhagem germinativa (Kondo e Ueda, 2013). Adicionalmente eles geraram nove outras linhagens transgênicas de drosófila, expressando gRNAs contra diferentes genes, todos sob controle do promotor constitutivo U6.

O cruzamento de cada uma dessas linhagens com *DmCas9* gerou proles (F₁) correspondentes, apresentando dois transgenes (Cas9 e um sgRNA). Apesar de o sgRNA ser expresso em todos os tecidos durante todo o desenvolvimento, Cas9 é expressa apenas na linhagem germinativa, portanto, a reconstituição do complexo Cas9/sgRNA ocorre apenas nestas células. Assim, os autores observaram uma taxa de edição genética média de 60% na linhagem germinativa desses animais. Por sua vez, os fenótipos derivados destes eventos são observados na F₂.

Exatamente no mesmo mês, outro grupo de pesquisadores publicou uma linhagem de drosófila praticamente idêntica. Contudo, neste trabalho, o sgRNA foi injetado em embriões *DmCas9*, permitindo a edição genética nos estágios iniciais do desenvolvimento (Ren et al., 2013). Dentro de poucos meses, vários outros trabalhos muito semelhantes foram publicados (Sebo et al., 2014; Gratz et al., 2014). Ainda em 2014, foram publicados os dados referentes às primeiras linhagens de drosófila transgênica expressando Cas9 de maneira ubíqua (*i.e.*, em todos os tecidos) (Port et al., 2014) e de expressão condicional (Xue et al., 2014).

Zebrafish Cas9

O pequeno tamanho, rápido desenvolvimento e baixo custo de manutenção tornaram o peixe paulistinha (*Dario rerio*) uma das principais espécies-modelo de pequenos vertebrados. Yin e colaboradores (2015) desenvolveram diversas linhagens transgênicas desta espécie para Cas9 (**DrCas9**). Duas delas apresentam expressão ubíqua de Cas9, sob controle dos promotores *ubiquitin* ou *beta2-actin*; uma linhagem com expressão tecido-específica (fígado), tendo Cas9 sob controle do promotor *fabp10* e uma linhagem para expressão condicional (*heat shock*), utilizando o promotor (*HotCre*).

Linhagens *DrCas9* foram cruzadas com linhagens transgênicas expressando sgRNAs sob o controle do promotor U6, gerando animais duplo transgênicos aptos para a

edição. Alternativamente, indivíduos adultos *DrCas9* foram injetados com gRNAs. Além de demonstrarem o sucesso de edição genética em todos estes cenários, os autores adicionalmente reportaram a realização de multiplex CRISPR, modificando vários genes simultaneamente em *zebrafish* (Yin et al., 2015).

***C. elegans* Cas9**

O nematoide de vida livre *C. elegans* apresenta um corpo translúcido com exatamente 959 células (~1 mm de comprimento), breve ciclo de vida (~72h de ovo a adulto) e também baixo custo de manutenção. Curiosamente, *C. elegans* se alimenta de bactérias (bacterióvoro) e, pensando nessa característica tão frívola, pesquisadores desenvolveram um protocolo revolucionário. Em 1998, uma dupla norte-americana demonstrou que é possível silenciar genes em *C. elegans* por meio da alimentação, *i.e.*, fornecendo bactérias transgênicas expressando dsRNAs (Timmons e Fire, 1998). A partir de então, os protocolos de RNAi, que fora desenvolvida neste mesmo ano, não precisariam mais ser realizados pelo laborioso e difícil processo de microinjeção de dsRNAs nas gônadas deste nematoide.

Mais recentemente, Liu e colaboradores (2014) apresentaram linhagens transgênicas de *C. elegans* com Cas9 sob controle de promotores para expressão ubíqua (*dpy-30*), tecido-específica (*ges-1*, predominantemente expresso no intestino; *unc-119* em neurônios; *unc-54*, músculos e *dpy-7/dpy-5*, hipoderme) ou indutível (*hsp-16.2*). Tendo em mente o trabalho de Timmons e Fire, eles testaram se CRISPR poderia ser desencadeada nestas linhagens pela simples alimentação com bactérias expressando sgRNAs. Para a agradável surpresa da comunidade que trabalha com esta espécie, a resposta é sim (Liu et al., 2014).

Camundongo Cas9

Visando à construção de um modelo murinho para o câncer, um grupo de pesquisadores do MIT anunciou em 2014 a criação de uma linhagem transgênica de camundongos expressando Cas9 de maneira constitutiva (ou apenas em certos tipos de neurônios). Estes animais foram gerados por um sistema mais sofisticado, de recombinação condicional (Cre-loxP) e, de acordo com os autores, todos os animais transgênicos foram viáveis, saudáveis e férteis. Inicialmente, os pesquisadores coletaram células dendríticas primárias (sistema imunológico) e conduziram experimentos de edição genética *ex vivo* com sucesso, pela simples infecção destas células com vetores lentivirais carregando sgRNAs.

Em seguida, o grupo administrou gRNAs *in vivo* por meio de lentivírus, vírus adenoassociados ou nanopartículas, obtendo sucesso na alteração de *KRAS*, *p53* e *LKB1*, genes cujas mutações estão associadas ao desenvolvimento de adenocarcinoma de pulmão. Assim, por meio destas estratégias, os autores conseguiram construir um modelo experimental que apresentasse determinadas mutações genéticas somáticas e, conseqüentemente, recapitulasse fenótipos (surgimento de tumores) em indivíduos adultos. Tal modelo possibilitará uma melhor compreensão deste tipo de câncer e, também, a realização de testes para o desenvolvimento de drogas.

Outras espécies Cas9

Há diversos trabalhos relatando a produção de plantas/animais duplo transgênicos, para a expressão de Cas9 e determinado sgRNA, visando à edição genética em um loco

endógeno (Ikeda et al., 2016). Contudo, estes organismos são triplamente modificados e não “puramente Cas9”. Devido ao elevado ritmo de pesquisas no presente, torna-se muito difícil acompanhar plenamente todos os desenvolvimentos na área de CRISPR. Portanto, é possível que já existam outras espécies de leveduras, plantas e animais transgênicos para Cas9.

Curiosamente, há casos em que plantas duplo transgênicas são produzidas para se obter determinada edição genética e, após isso, por meio de cruzamentos controlados, os pesquisadores conseguem chegar a indivíduos contendo apenas a edição genética, sem os transgenes para Cas9 ou para o sgRNA (Zhou et al., 2014), exatamente o oposto do cenário vislumbrado nesta seção.

Riscos da expressão heteróloga de Cas9

Por fim, é essencial lembrarmos que Cas9 é uma nuclease bacteriana, não havendo homólogos em leveduras, plantas ou animais. Portanto, as células eucarióticas não estão preparadas para “lidar” com uma Cas9 (assim qualquer outro transgene). Uma vez dentro da célula, Cas9 poderia interagir de maneira imprevista com transcritos celulares cuja estrutura se assemelhe a um sgRNA, levando assim a edições genéticas inesperadas. Este potencial risco é ainda maior em espécies transgênicas expressando constitutivamente Cas9. Análises detalhadas das implicações deste potencial risco, capaz de gerar alterações fenotípicas de fundo, foram abordadas por Souza e Pereira (2014).

Conclusões

A criação de diferentes espécies transgênicas expressando Cas9 em diversos padrões espaço-temporais (constitutivo, ubíquo, tecido-específico ou condicional) simplificou ainda mais a poderosa ferramenta de edição genética CRISPR. Estas linhagens podem ser utilizadas de diversas formas: (i) administração exclusiva do gRNA (transfecção, infecção, eletroporação, etc.), (ii) cruzamento com outras linhagens transgênicas expressando gRNAs, gerando prole duplo transgênica ou (iii) via alimentação. Em breve, o repertório de espécies e linhagens celulares transgênicas puramente para Cas9 deve se ampliar imensamente, mas investigações sobre possíveis interações espúrias entre esta nuclease e os transcriptomas destas células/organismos devem ser realizadas, via RIP (*RNA immunoprecipitation*) contra Cas9 por exemplo, para excluir tais possíveis interferências/edições inesperadas.

6. Conclusões do capítulo

No ramo da física, a observação de que o *universo está em expansão* foi em si notória (Hubble, 1929). A recente descoberta de que *essa expansão está em ritmo acelerado* foi algo igualmente surpreendente (Floyd, 1998). CRISPR em si pode ser interpretada como uma otimização de um processo de edição genética iniciada em 1985, com o *Gene Targeting* (vide cap. 5). Este processo de melhoria avançou lentamente em 30 anos, mas agora ele está em ritmo acelerado – como a própria expansão do universo.

Referências

- Albers, J. et al. A versatile modular vector system for rapid combinatorial mammalian genetics. *J Clin Invest.* 2015;125(4):1603-19.
- Aylon Y, Kupiec M. DSB repair: the yeast paradigm. *DNA Repair.* 2004;3:797-815.
- Basu S, Aryana A, Overcash JM, Samuela GH, Anderson MAE, Dahlemb TJ, Mylesa KM, Adelmana ZN. Silencing of end-joining repair for efficient sitespecific gene insertion after TALEN/CRISPR mutagenesis in *Aedes aegypti*. *Proc Natl Acad Sci USA.* 2015;112(13):4038-43.
- Beumer KJ, Trautman JK, Mukherjee K, Carroll D. Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3.* 2013;3:657-64.
- Buj, R. et al. A plasmid toolkit for cloning chimeric cDNAs encoding customized fusion proteins into any Gateway destination expression vector. *BMC Mol Biol.* 2013;14(1):18.
- Burma S, Chen BP, Chen DJ. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair.* 2006;5:1042-8.
- Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol.* 2015;33:543-8.
- Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013 fev.;339(6121):819-23.
- Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res.* 2013;2:130-43.
- Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 2013;47:433-55.
- Engler C, Marillonnet S. Golden Gate cloning. *Methods Mol Biol.* 2014;1116:119-31.
- Fausser F, Rotha N, Pachera M, Ilga G, Sánchez-Fernández R, Biesgenb C, et al. In planta gene targeting. *Proc Natl Acad Sci USA.* 2012;109(19):7535-40.
- Floyd E. Bloom. Breakthroughs 1998. *Science.* 1998;282(5397):2193.
- Fuller KK, Chen S, Loros JJ, Dunlap JC. Development of the CRISPR/Cas9 System for Targeted Gene Disruption in *Aspergillus fumigatus*. *Eukaryot Cell.* 2015 Nov;14(11):1073-80.
- Gantz VM, Bier E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science.* 2015;348(6233):442-4.
- Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc Natl Acad Sci USA.* 2015;112(49):E6736-43.
- Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics.* 2014 Apr;196(4):961-71.
- Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* 2013;31(9):827-32.
- Hubble E. A relation between distance and radial velocity among extra-galactic nebulae. *Proc Natl Acad Sci USA.* 1929;15(3):168-73.
- Ikedo T, Tanaka W, Mikami M, Endo M, Hirano HY. Generation of artificial drooping leaf mutants by CRISPR-Cas9 technology in rice. *Genes Genet Syst.* 2016;90(4):231-5.
- Isaacs AT, et al. Engineered resistance to *Plasmodium falciparum* development in transgenic *Anopheles stephensi*. *PLoS Pathog.* 2011;7(4):e1002017.
- Isaacs AT, et al. Transgenic *Anopheles stephensi* coexpressing single-chain antibodies resist *Plasmodium falciparum* development. *Proc Natl Acad Sci USA.* 2012;109(28):E1922-30.
- Jakočinas T, et al. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab Eng.* 2015;28:213-22.
- Jasin M, Rothstein R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol.* 2013;5:a012740.
- Jiang Y, et al. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol.* 2015;81(7):2506-14.
- Kabadi AM, et al. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* 2014;42(19):e147-147.
- Kondo S, Ueda R. Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. *Genetics.* 2013 Nov;195(3):715-21. <http://dx.doi.org/10.1534/genetics.113.156737>.
- Konermann S, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* 2014 dez.;517(7536):583-8.
- Li F, Patra KP, Vinetz JM. An anti-Chitinase malaria transmission-blocking single-chain antibody as an effector molecule for creating a *Plasmodium falciparum*-refractory mosquito. *J Infect Dis.* 2005;192(5):878-87.
- Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem.* 2010;79:181-211.
- Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife.* 2014;3:e04766.
- Liu P, Long L, Xiong K, Yu B, Chang N, Xiong JW, et al. Heritable/conditional genome editing in *C. elegans* using a CRISPR-Cas9 feeding system. *Cell Res.* 2014 Jul;24(7):886-9.
- Maddalo D, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature.* 2014;516(7531):423-7.
- Maeder ML, Gersbach CA. Genome Editing Technologies for Gene and Cell Therapy. *Mol Ther.* 2016;1-17. <http://dx.doi.org/10.1038/mt.2016.10>.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013;339:823-6.
- Maruyama T, Dougan SK, Truttmann M, Bilate AM, Ingram JR, Ploegh HL. Inhibition of non-homologous end joining increases the efficiency of CRISPR/Cas9-mediated precise (TM: inserted) genome editing. *Nat Biotechnol.* 2015;33(5):538-42.
- Mefferd AL, et al. Expression of CRISPR/Cas single guide RNAs using small tRNA promoters. *RNA.* 2015;99:rna.051631.115.

- Nissim L, et al. Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol Cell*. 2014;54(4):698-710.
- Niu Y, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*. 2014;156(4):836-43.
- Nødvig CS, et al. A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLoS One*. 2015;10(7):e0133085.
- Ochi T, Blackford AN, Coates J, Jhujh S, Mehmood S, Tamura N, et al. DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science*. 2015;347:185-8.
- Platt RJJ, et al. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell*. 2014;159(2):440-55.
- Port F, Chen HM, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc Natl Acad Sci USA*. 2014 Jul;111(29):E2967-76.
- Puchta H, Fauser F. Gene targeting in plants: 25 years later. *Int J Dev Biol*. 2013;57:629-37.
- Ran FA, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-308.
- Ren X, Housden BE, Hu Y, Roesel C, Lin S, et al. Optimized gene editing technology for *Drosophila melanogaster* using germ line-specific Cas9. *Proc Natl Acad Sci USA*. 2013 Nov;110(47):19012-7.
- Robert F, Barbeau M, Éthier S, Dostie J, Pelletier J. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med*. 2015;7:93.
- Ryan OW, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. *eLife*. 2014;3:1-15.
- Sakuma T, et al. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep*. 2014;4:5400.
- Schimpl S, Fauser F, Puchta H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny. *Plant J*. 2014;80:1139-50.
- Schimpl S, Puchta H. Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. *Plant Methods*. 2016;12:8.
- Scott WG, Horan LH, Martick M. The Hammerhead Ribozyme. *Prog Mol Biol Transl Sci*. 2013;120:1-23.
- Sebo XL, Lee HB, Peng Y, Guo Y. A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. *Fly*. 2014;8(1):52-7.
- Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng*. 2008;2(1):5.
- Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. *Genetics*. 2015;199:1-15.
- Sinkins SP, Gould F. Gene drive systems for insect disease vectors. *Nat Rev Genet*. 2006;7(6):427-35.
- Souza TAJ, Pereira TC. Cell. 2014. Available from: [http://www.cell.com/cell/comments/S0092-8674\(14\)01163-5](http://www.cell.com/cell/comments/S0092-8674(14)01163-5).
- Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, et al. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell*. 2012;151:1474-87.
- Srivastava M, Raghavan SC. DNA double-strand break repair inhibitors as cancer therapeutics. *Chem Biol*. 2015;22:17-29.
- Stone D, Niyonzima N, Jerome KR. Genome editing and the next generation of antiviral therapy. *Hum Genet*. 2016;1-12. <http://dx.doi.org/10.1007/s00439-016-1686-2>.
- Swiech L, et al. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol*. 2015;33(1):102-6.
- Timmons L, Fire A. Specific interference by ingested dsRNA. *Nature*. 1998 Oct;395(6705):854.
- Tsai SQ, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32(6):569-76.
- Unckless RL, Messer PW, Connallon T, Clark AG. Modeling the Manipulation of Natural Populations by the Mutagenic Chain Reaction. *Genetics*. 2015;201(2):425-31.
- Vartak SV, Raghavan SC. Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing. *FEBS J*. 2015;282(22):4289-94.
- Wang H, et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell*. 2013;153(4):910-8.
- Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;153:910-8.
- Wang X, et al. One-step generation of triple gene-targeted pigs using CRISPR/Cas9 system. *Sci Rep*. 2016;6:20620.
- Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. *Annu Rev Genet*. 2006;40:363-83.
- Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc Natl Acad Sci USA*. 2015;112(11):3570-5.
- Xing H-L, et al. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol*. 2014;14(1):327.
- Xue Z, Ren M, Wu M, Dai J, Rong YS, Gao G. Efficient gene knock-out and knock-in with transgenic Cas9 in *Drosophila*. *G3*. 2014 Mar;4(5):925-9.
- Xue Z, Wu M, Wen K, Ren M, Long L, Zhang X, et al. CRISPR/Cas9 mediates efficient conditional mutagenesis in *Drosophila*. *G3*. 2014 Sep;4(11):2167-73.
- Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;154:1370-9.
- Yin L, Maddison LA, Li M, Kara N, LaFave MC, Varshney GK, et al. Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs. *Genetics*. 2015 Jun;200(2):431-41.
- Yoshioka S, et al. Development of a mono-promoter-driven CRISPR/Cas9 system in mammalian cells. *Sci Rep*. 2015;5:18341.
- Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y, et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell*. 2015;16:142-7.
- Zhang WW, Matlashewski G. CRISPR-Cas9-mediated genome editing in *Leishmania donovani*. *MBio*. 2015;6(4):1-14.
- Zhou H, Liu B, Weeks DP, Spalding MH, Yang B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res*. 2014;42(17):10903-14.

Questões éticas, legais, ambientais e de pioneirismo

Capítulo 11

Nilza Maria Diniz^{1,a}, Andréa Almeida Carneiro^{2,b}, Newton Portilho Carneiro^{2,b},
Camila Tenorio França^{3,4,c}, Maria Betânia Melo de Oliveira^{5,c}
e Giulliana Augusta Rangel Gonçalves^{6,d}

¹Dept.º de Ciências Biológicas, CCB, UEL, PR - Brasil

²Núcleo de Biologia Aplicada, Embrapa Milho e Sorgo, MG - Brasil

³Population Health and Immunity Division, Walter and Eliza Hall Institute of Medical Research,
Melbourne - Austrália

⁴Dept. of Medical Biology, University of Melbourne, Melbourne - Austrália

⁵Dept.º de Bioquímica, CCB, UFPE, PE - Brasil

⁶Instituto Israelita de Ensino e Pesquisa, SP - Brasil

^aAutoria da seção: 2

^bAutoria da seção: 3

^cAutoria da seção: 4

^dAutoria da seção: 5

Estruturação do capítulo

1. Introdução
2. Éticas
 - 2.1 Bioética e genética
 - 2.2 A edição de embriões humanos
 - 2.3 A questão moral
3. Legais
 - 3.1 A regulamentação da edição genética
 - 3.2 O quadro regulatório mundial
4. Ambientais
 - 4.1 Introdução
 - 4.2 Potenciais aplicações da técnica MCR
 - 4.3 Cuidados no uso da técnica e regulamentações vigentes
 - 4.4 Riscos ambientais da dispersão de genes derivados de MCR
 - 4.5 Controvérsias sobre a técnica MCR e sua segurança
 - 4.6 Conclusão
5. Pioneirismo
 - 5.1 As patentes e o mercado da biotecnologia
 - 5.2 A disputa Doudna-Charpentier versus Zhang
6. Conclusões

1. Introdução

Os desdobramentos de CRISPR são tão extensos que ultrapassam o campo técnico, atingindo as esferas sociais (a bioética), políticas (regulamentação de OGMs), ecológicas (impacto ambiental) e econômicas (patentes). Neste capítulo abordaremos um pouco sobre o *status quo* da CRISPR nestes outros contextos.

2. Éticas

2.1 Bioética e genética

O neologismo “Bioética”, cunhado por Potter na década de 1970 a partir das palavras do grego *Bios* (vida) e *Ethos* (ética), com o sentido de “ciência da sobrevivência humana”, trouxe a ideia de uma nova ciência que

... combina o trabalho dos humanistas e cientistas, cujos objetivos são sabedoria e conhecimento. A sabedoria é definida como o conhecimento de como usar o conhecimento para o bem social. A busca de sabedoria tem uma nova orientação porque a sobrevivência do homem está em jogo. Os valores éticos devem ser testados em termos de futuro e não podem ser divorciados dos fatos biológicos. (Potter, 1971)

Em 19 de abril 1971, quando a revista norte-americana *Times* publicou o artigo “*Man into superman: the promise and peril of the new genetics*” (Homem em Super-Homem: a promessa e perigo da nova genética), o neologismo “Bioética” apareceu na mídia pela primeira vez, segundo Leo Pessini (2013).

A bioética e a genética estiveram sempre relacionadas desde a origem de uma e expansão da outra. Naquele momento, a possibilidade de manipulação genética ainda era algo para o “Futuro”. Iremos nesta seção discutir o “Futuro” previsto naquele momento, ou o nosso “Presente”, o qual nem os mais criativos poderiam ter imaginado a amplitude ou precisão das manipulações genéticas, ou seja, a edição de embriões humanos pela utilização da técnica de CRISPR; e a partir daí pensar o futuro uso desta técnica.

A questão a ser respondida seria: “De que maneira a utilização desta técnica poderia comprometer o futuro e diminuir as chances de sobrevivência humana?”. Os fatos biológicos estão dados e há necessidade de reflexão sobre futuro uso desta técnica de edição de material genético de embriões humanos. Segundo Potter, “*Ações que diminuem as chances de sobrevivência humana são imorais e devem ser julgadas em termos do conhecimento disponível e no monitoramento de ‘parâmetros de sobrevivência’ que são escolhidos pelos cientistas e humanistas...*” (Potter, 1971) Então, de que maneira iremos nos preparar para esta nova fase da genética sem comprometer nossa existência?

O “super-homem” apresentado no artigo da *Times* referia-se ao “*Übermensch*” de Friedrich Nietzsche (1998), ou seja, a meta que a humanidade pode dar para si mesma. Dito de outra maneira, “*Toda vida humana teria sentido caso ela melhorasse a cada nova geração de seres humanos*”.

Ao escrever sobre a obra de Nietzsche, Warburton afirma que

Em Assim falou Zaratustra (1883-1892), Nietzsche escreveu sobre o Übermensch ou “Sobre-homem”. O termo descreve uma pessoa imaginada no futuro que não está presa aos códigos morais convencionais, mas vai além deles, criando novos valores. Talvez influenciado pelo próprio entendimento da teoria da evolução de Charles Darwin, Nietzsche tenha visto o Übermensch como o próximo passo no desenvolvimento da humanidade.

A técnica de “edição” do genoma chamada CRISPR-Cas9 torna possível para os cientistas inserir, remover ou corrigir o DNA de forma simples e eficiente. Ela mantém a perspectiva de tratar ou mesmo curar certas enfermidades, tais como doenças falciformes, fibrose cística e alguns tipos de câncer. Mas a edição de linhagens germinativas pode também tornar mais fácil aos cientistas que trabalham com embriões humanos, óvulos e espermatozoides, alterações ao DNA, tais como as que determinam cor dos olhos de um bebê (UNESCO¹).

A biotecnociência apresenta as condições para a manipulação dos genomas (*conhecimento*) e a bioética tem o desafio de apresentar possíveis usos das biotecnologias desenvolvidas na genética (*conhecimento de como usar o conhecimento*) dessas novas biotecnologias. Vale destacar que esta sabedoria de como usar o conhecimento deve ser exercida tanto por cientistas como humanistas. Desta forma, faz-se necessário que todos tenham o mínimo de instrumentos de reflexão sobre os impactos das biotecnologias para a humanidade e para o planeta. Nem tudo que PODE tecnicamente ser executado DEVE ser executado.

O termo biotecnociência foi definido ainda nos anos 1990 por Schramm. Segundo ele

“... é preciso distinguir duas formas de tecnologia: a tecnologia convencional, que pertence ao campo da tecnociência, e a biotecnologia, pertencente ao campo daquela que chamaremos biotecnociência. A distinção entre as duas está contida no radical bio, que indica ao mesmo tempo uma inovação e o retorno a uma tradição pré-moderna, isto é, à metáfora “natural”. Com efeito, contrariamente aos artefatos da tecnociência clássica – que são verdadeiros artefatos inventados sem em princípio nenhum modelo pré-existente – as tecnologias de segunda geração operam tendo um “modelo” prévio, encontrado, de regra, nos próprios processos naturais, ou em algum outro “exemplar” pré-existente.

Seria recomendável que aqueles que estão mergulhados na esfera da ciência, confortável e segura em sua execução de conclusões, em alguns momentos se distanciassem desta esfera e percebessem o “dever” agir com prudência e responsabilidade, e não apenas o “poder” executar tecnicamente como máxima da ação.

2.2 A edição de embriões humanos

Casos Chinês e Britânico

CRISPR foi utilizada em embriões humanos pela primeira vez em 2015. Pesquisadores chineses utilizaram a técnica em embriões “não viáveis” (Liang, 2015), fato que causou

¹ UNESCO, UNESCO panel of experts calls for ban on “editing” of human DNA to avoid unethical tampering with hereditary traits. Disponível em < http://www.unesco.org/new/en/media-services/single-view/news/unesco_panel_of_experts_calls_for_ban_on_editing_of_human_dna_to_avoid_unethical_tampering_with_hereditary_traits/#.VzxLttQrJH0 > Acesso em: 10 maio 2016.

grande repercussão e polêmica. O objetivo do grupo da *Sun Yat-sen University* era modificar o gene associado à β -talassemia. Os embriões utilizados no experimento eram produtos de fecundação de um óvulo com dois espermatozoides, obtidos em uma clínica de fertilização humana. Apesar de não haver possibilidade de desenvolvimento de um ser humano, a controvérsia foi grande.

No Reino Unido, foi aprovado pelo “*Fertilisation and Embryology Authority*” (HFEA), em 1º de fevereiro de 2016, o projeto para edição de embrião humano saudável, proposto por Kathy Niakan do “*Francis Crick Institute*” (Callaway, 2016). Neste caso, contudo, o objetivo da equipe será estudar o desenvolvimento inicial, e planeja alterar os genes que são ativos durante os primeiros dias após a fertilização.

A proposta do grupo seria bloquear a atividade de um gene ‘regulador mestre’ chamado *OCT4*, também conhecido como *POU5F1*, que está ativo em células que vão formar o feto (e não nas células que formarão a placenta). A equipe da Dr.^a Niakan planeja encerrar os experimentos após uma semana da fecundação, quando o óvulo fertilizado atingir o estágio de blastocisto de desenvolvimento, contendo até 256 células. Neste momento, os embriões serão destruídos. Os fatos estão colocados, já há manipulação de embriões humanos em curso.

Moratória

A controvérsia gerada pela primeira manipulação de embrião humano via CRISPR levou alguns pesquisadores a proporem uma moratória (Cyranoski e Reardon, 2015). Houve no passado uma moratória proposta pela “Reunião de Asilomar”, na Califórnia (Estados Unidos), envolvendo 140 cientistas (entre eles biólogos e médicos), advogados e 16 jornalistas liderados por Paul Berg e David Baltimore, para discussão da manipulação de DNA em 1974 (Goldim, 2015).

Segundo Goldim (2015), o documento resultado da Conferência Internacional sobre *DNA recombinante* seria um precursor do princípio da precaução. Este princípio afirma que qualquer risco de dano grave ou irreversível requer a tomada de medidas para prevenir a sua ocorrência.

O princípio da precaução é uma abordagem importante para este problema, porque CRISPR cumpre os critérios de riscos potencialmente graves e imprevisíveis, se não forem implementadas medidas para prevenir esses riscos antes que haja dano real. Embora seja muito mais arriscado do que a recombinação do DNA, o que levou à Conferência de Asilomar, ou o Projeto Genoma Humano, que propôs uma reflexão sobre as implicações Éticas, Legais e Sociais suscitadas pelo projeto criando um comitê específico e permanente que na origem era chamado de ELSI (*Ethics, legal and Social Implications*) associado com mapeamento do genoma humano, a técnica de CRISPR tem sido cada vez mais utilizada, mas sem reflexão adequada sobre as suas muitas implicações. CRISPR é, portanto, um dos desafios mais importantes apresentados pela ética e pesquisa genética no presente (Goldim, 2015).

Contudo, alguns pesquisadores contestaram a necessidade de moratória para a utilização da técnica de CRISPR em embriões humanos (Cyranoski e Reardon, 2015), como John Harris da *Manchester University* (Reino Unido), por exemplo, que afirmou que não seria “*pior do que o que acontece na FIV (fertilização in vitro) o tempo todo, na qual os embriões não viáveis são descartados. Eu não vejo qualquer justificativa para uma moratória sobre as pesquisas*”. Outros, como Tetsuya Ishii da *University of Hokkaido* (Japão), acreditam que por ter sido

utilizado embrião criado para fertilização *in vitro*, e não para pesquisa, já se teria superado muitos dos obstáculos éticos que outros países enfrentariam (Cyranoski e Reardon, 2015). Ou ainda, para George Daley da *Harvard Medical School*, a edição de genes também poderia ser usada para manipular mutações relacionadas com alguma doença específica em um embrião, o qual poderia então ser utilizado para produzir células-tronco embrionárias que poderiam funcionar como modelos para o ensaio de drogas e outras intervenções para a doença (Cyranoski e Reardon, 2015).

O princípio da precaução já estabelecido na década de 1970 e aprofundado nas últimas décadas poderia ser o princípio norteador das ações em relação à utilização da técnica de CRISPR em seres humanos.

As medidas para prevenir riscos de dano grave ou irreversível podem ser recomendadas em nível internacional com os documentos da UNESCO. A “Declaração Universal de Bioética e Direitos Humanos” incorporou tanto a “Declaração Universal do Genoma Humano e Direitos Humanos”, adotada em 1997, pela Conferência Geral da UNESCO, quanto a “Declaração Internacional de Dados Genéticos Humanos”, adotada em 2003 (Cruz et al., 2010), as quais trazem a obrigação de respeito à dignidade humana, assumindo que

O genoma humano constitui a base da unidade fundamental de todos os membros da família humana bem como de sua inerente dignidade e diversidade. Em um sentido simbólico, é o patrimônio da humanidade. (Artigo 1- Declaração Universal do Genoma Humano e Direitos Humanos) (UNITED NATIONS, 1997).

Em outubro de 2015, alguns meses depois da publicação do trabalho dos chineses, a UNESCO publicou documento de atualização sobre manipulação genética². Segundo este relatório³ da UNESCO em seu item “II. 2.5. *Benefits versus burdens*” (Benefícios e encargos), a “*Pesquisa só deve ser permitida para o benefício da humanidade, e seus resultados não devem ser utilizados para fins não pacíficos como o bioterrorismo ou fins que violem a dignidade humana, como a Eugenia.*” Também afirma que

Devido às incertezas sobre o efeito da modificação da linhagem germinativa sobre as gerações futuras, tais intervenções têm sido fortemente desencorajadas ou legalmente proibidas em muitos países. Há casos excepcionais em que as intervenções sobre o genoma podem ser realizadas apenas para fins preventivos, de diagnóstico ou terapêuticos, e não apenas se o objetivo é introduzir qualquer modificação hereditária no genoma.

Tendo sido estabelecido o princípio da precaução há muito tempo, e com as medidas internacionais de reiterar as obrigações entre os Estados, não se faz necessário uma nova moratória como a proposta no encontro de NAPA na Califórnia (Cyranoski, 2015).

² UNESCOPRES, UNESCO panel of experts calls for ban on “editing” of human DNA to avoid unethical tampering with hereditary traits. Disponível em < http://www.unesco.org/new/en/media-services/single-view/news/unesco_panel_of_experts_calls_for_ban_on_editing_of_human_dna_to_avoid_unethical_tampering_with_hereditary_traits/#.VzxLttQrJH0 > Acesso em: 10 maio 2016.

³ International Bioethics Committee (IBC) Report of the IBC on Updating Its Reflection on the Human Genome and Human Rights. Disponível em <<http://unesdoc.unesco.org/images/0023/002332/233258E.pdf>> Acesso em: 10 maio 2016.

2.3 A questão moral

Se nem tudo que PODE tecnicamente ser executado DEVE efetivamente ser executado, alguns valores são colocados para nortear o “DEVER fazer” frente à possibilidade de utilização de CRISPR, como, por exemplo, erradicar doenças, em contraponto com o valor intrínseco do genoma como patrimônio da humanidade; o valor da ciência e o valor do ser humano; entre outros.

Poderíamos pensar o agir do cientista em qualquer questão moral originada da utilização desta técnica de edição de embriões humanos em uma perspectiva de reflexão proposta por Siqueira e Diniz (2003), na qual se considera que

“Bom seria, enfim, que os cientistas considerassem, no agir cotidiano, os alertas de Jaspers sobre o uso inadequado da tecnociência como instrumento de domínio, percebessem o apelo de Jonas a humildade e ponderação ante aos avanços científicos, acolhessem o chamado de Morin para construir uma sociedade baseada na compreensão, na solidariedade e na compaixão. Finalmente que estivessem atentos a urgência universal de uma conversão existencial que fizesse justiça aos direitos do outro, sobretudo aos mais fracos.

Se não podemos alcançar a complexidade dos fatos e as consequências do agir frente às novas tecnologias, proponho que uma sabedoria mínima se faz necessária, ou seja, agir com prudência (Aristóteles, 2001).

Há cento e cinquenta anos, Mendel propôs a existência de fatores que determinariam as características hereditárias; hoje queremos determinar as características, ou melhor, determinar os fatores que cada indivíduo carregará em seu genoma.

Voltando ao ponto de partida de nossa reflexão, o homem define sua própria existência para Nietzsche, mas, a meu ver, não haveria respaldo moral para definir sua porção biológica e, quem sabe, comprometer sua própria existência com a manipulação de embriões via CRISPR.

Deveremos, portanto, numa visão aristotélica, procurar a justa medida para a utilização desta técnica sem impedir os benefícios que ela venha a trazer para cura de doenças e, ao mesmo tempo, evitar abusos em relação à integridade da nossa própria espécie.

3. Legais

3.1 A regulamentação da edição genética

A utilização da biotecnologia para a produção de organismos geneticamente modificados (OGM) surgiu como um suporte ao melhoramento de espécies, uma vez que este é limitado pela disponibilidade de alelos benéficos na natureza (Bortesi e Fischer, 2015). Mesmo depois de vinte anos de comercialização das primeiras culturas biotecnológicas e o plantio aproximado de dois bilhões de hectares em 28 países (ISAAA, 2016), essas culturas ainda são uma fonte inesgotável de controvérsia pública. Parte desta polêmica está relacionada com a inserção aleatória de genes no produto final e o potencial aparecimento de efeitos indesejáveis (Pauwels et al., 2014). O questionamento público da biossegurança de organismos transgênicos tem

causado aumento dos encargos regulatórios, atrasos em pesquisas, restrição do número de culturas comercializadas e de empresas capazes de desenvolver este tipo de produto (Wolt et al., 2016).

Atualmente, a comunidade científica está presenciando um novo avanço tecnológico conhecido como edição de genoma. Em oposição às técnicas iniciais utilizadas para a produção de OGMs, estas novas tecnologias utilizam complexos contendo nucleases capazes de clivar o DNA em locais específicos. Durante o processo de reparo da clivagem podem ocorrer uma inserção ou deleção, levando à mutação do loco desejado.

A história foi marcada pelo desenvolvimento de três grandes tecnologias de edição de genoma: ZFN, TALEN e CRISPR (vide cap. 5). Em oposição aos sistemas de edição de genoma anteriores, CRISPR é mais simples, eficiente e versátil. Uma vez realizada a modificação pretendida no genoma, o complexo Cas/gRNA pode ser removido utilizando diferentes estratégias moleculares (Curtin et al., 2012; Wook et al., 2015) e a região editada é transmitida às gerações subsequentes (Curtin et al., 2012).

Embora as tecnologias de edição do genoma sejam bastante estridentes no reconhecimento de regiões específicas do DNA, CRISPR possui a limitação de induzir possíveis mutações fora do alvo, conhecidas como “*off-targets*” (Shan et al., 2013; Mahfouz et al., 2014). De acordo com Araki e Ishii (2015), mutações não específicas podem resultar em silenciamento ou perda de função gênica, também podem levar à imunogenicidade ou toxicidade nos produtos alimentares devido à mudança de aminoácidos dentro de uma proteína. Além disso, tem sido especulado que o cultivo de culturas com mutações não precisas pode afetar um ecossistema como resultado de cruzamentos (Prasad et al., 2012). Apesar de o problema não estar completamente solucionado, pesquisas têm sido realizadas para reduzir a frequência de indução de mutações não específicas causadas por esta técnica (Slaymaker et al., 2016; Kleinstiver et al., 2016).

Semelhante ao que ocorreu com a transformação genética e a produção de OGMs, as modernas técnicas de edição de genoma também estão suscitando questões regulatórias. Ferramentas de edição do genoma podem gerar organismos que são extremamente similares ou idênticos aos existentes na natureza, ou seja, uma variação genética natural, criando limites indistinguíveis entre uma modificação que ocorreu naturalmente ou uma que foi induzida por nucleases sítio-específicas (Araki et al., 2014). Portanto, uma resposta dos reguladores à aplicação destas novas técnicas é urgente para o aceite dos novos organismos gerados (Araki e Ishii, 2015).

3.2 O quadro regulatório mundial

As tecnologias de edição de genomas através de nucleases sítio-direcionadas (NSD) como a CRISPR podem ser utilizadas para a introdução de: (i) mutações aleatórias sítio-específicas (mudanças de apenas um par de bases, pequenas deleções ou inserções, conhecidas como *indels*) induzidas após a junção das extremidades clivadas do DNA (NSD1); (ii) mutações pontuais sítio-específicas desejadas (substituições específicas de um ou poucos nucleotídeos ou pequenas deleções ou inserções, conhecidas como *substituições alélicas*) induzidas após o reparo por recombinação homóloga da fita de DNA clivada. Neste processo, um DNA homólogo é adicionado e compete com as cromátides-irmãs como um modelo de reparo, levando à substituição da sequência de nucleotídeos original (NSD2); e (iii) inserções

Tabela 1. Categorização das técnicas de edição de genoma através de nucleases sítio-direcionadas (NSD) e a tendência de regulamentação utilizando a legislação de OGMs.

Técnicas de edição de genomas	Modificação resultante	DNA exógeno no produto final	Regulamentação Transgênico / Não Transgênico
NSD1	Pequenos <i>indels</i> gerando mutações randômicas sítio-específicas sem recombinação homóloga	Ausente	Podem não sofrer regulamentação (exceto na União Europeia)
NSD2	Substituições específicas de um ou poucos nucleotídeos, com recombinação homóloga	Presente	Regulamentação com análise caso a caso
NSD3	Inserções sítio-direcionadas de transgenes	Presente	Regulamentação de maneira semelhante aos OGMs

sítio-direcionadas de transgenes (NSD3) por recombinação homóloga (EFSA, 2012; Wolt et al., 2016) (Tabela 1).

De acordo com o Protocolo de Biossegurança de Cartagena (<http://bch.cbd.int/protocol/text/>), um OGM é um organismo que possui uma combinação nova de material genético obtido através do uso da biotecnologia moderna. Esta definição alude que alguns organismos modificados por edição de genoma podem estar fora do âmbito das regulamentações atuais de OGM, porque esta técnica pode produzir segregantes que não possuem nenhuma inserção transgênica (Araki et al., 2014; Hartung e Schiemann, 2014; Voytas e Gao, 2014; Camacho et al., 2014). Os produtos gerados via tecnologia NSD1 são simples mutações pontuais e não podem ser discriminados das naturais, pois não carregam nenhum vestígio da técnica utilizada para produzi-los; em geral reguladores, não consideram estes organismos no mesmo contexto dos OGMs (Wolt et al., 2016). Já aqueles gerados via NSD3 possuem transgenes inseridos no genoma e serão regulados como OGMs, mas por causa da inserção sítio-direcionada do transgene poderá ser exigida, dependendo do quadro regulatório do país, uma menor quantidade de dados para a caracterização dos riscos quando comparado com um transgênico convencional (Wolt et al., 2016). Já a regulamentação para produtos enquadrados em NSD2 não é clara, pois depende da natureza e extensão da edição realizada, podendo estes ser ou não contidos na regulamentação de OGMs (Lusser e Davies, 2013).

A necessidade de protocolos regulatórios definindo como lidar com os organismos provenientes das técnicas de NSD já vem sendo discutida há algum tempo em diferentes países. Argentina, Austrália, Brasil, Nova Zelândia, União Europeia, dentre outros, possuem sistemas de regulamentação específicos para avaliar seus organismos GM (Podevin et al., 2012) e, no momento, estão ponderando se os organismos gerados através da edição de genoma por NSD se enquadram nesta regulamentação. Na Argentina em 2011, um seminário regulatório preliminar para analisar as novas tecnologias aplicadas ao melhoramento de organismos definiu que apenas produtos baseados na edição de genoma a partir de técnicas NSD1 poderiam não ser incluídos na regulamentação de OGMs, NSD3 seriam regulamentados e NSD2 seriam analisados caso a caso (Burachik, 2011).

Austrália e Nova Zelândia, semelhantes à Argentina, concordam que produtos NSD3 devem estar sob regulamentação. Em 2013, o Comitê de Proteção Ambiental (EPA) da Nova

Zelândia declarou que plantas NSD1 não seriam regulamentadas como OGMs, mas em 2014 esta decisão foi nulificada.

A legislação da União Europeia considera que um organismo transgênico é todo aquele no qual o material genético foi alterado de uma maneira que não ocorre naturalmente por cruzamento e/ou recombinação. Em 2012, um comunicado científico da Autoridade Europeia de Segurança Alimentar (EFSA) concluiu que alimentos e rações derivados de produtos gerados pela técnica NSD3 poderiam causar menos perigo do que os transgênicos convencionais, devido ao fato de a inserção gênica ter sido realizada em uma região predefinida do genoma. Estes produtos deveriam ser enquadrados dentro da legislação da União Europeia para OGMs. A Comissão Europeia ainda não emitiu parecer algum relacionado às técnicas NSD1 e NSD2 (Jones, 2015).

No Brasil, grupos de pesquisadores e legisladores têm-se reunido para definir se os produtos gerados pelas modernas técnicas de edição de genoma deverão se enquadrar ou não na legislação de transgênicos. Entretanto, até o presente momento nenhum comunicado oficial sobre este assunto foi emitido pela Comissão Técnica de Biossegurança (CTNBio), órgão ligado ao Ministério da Ciência, Tecnologia e Inovação que assessora o Governo Federal em questões relativas à biossegurança de OGMs.

O Canadá e os EUA optaram por regular todas as plantas ou produtos com novas características desenvolvidas por meio de engenharia genética ou quaisquer outras técnicas de melhoramento de plantas com o mesmo, já existente, sistema de regulamentação. O foco dos sistemas de regulação é sobre os riscos inerentes ao novo produto, em vez do método de produção (Podevin et al., 2012; Smyth e McHughen, 2008; Wolt et al., 2016).

Nos Estados Unidos, os reguladores indicam que abordagens sítio-direcionadas que resultam em exclusões específicas de nucleotídeos endógenos (NSD1) não seriam reguladas pelo Departamento de Agricultura Americano (USDA). Da mesma maneira, abordagens iniciadas com transgenia, nas quais passos intermediários selecionam para a ausência dos elementos transgênicos, não seriam reguladas pelo USDA. No entanto, métodos de mutação sítio-direcionadas envolvendo inserções ou substituições de oligonucleotídeos deverão ser analisados caso a caso (Camacho et al., 2014).

Eventos gerados por tecnologias baseadas na edição de genomas, como uma linhagem de milho contendo baixa concentração de fitato, gerada utilizando a tecnologia ZNF, foram classificados como fora do escopo da regulamentação de OGMs pelo USDA (Waltz, 2012). Além disso, recentemente o USDA emitiu um parecer que não irá regulamentar o cogumelo geneticamente modificado, no qual se utilizou a tecnologia CRISPR para evitar o escurecimento do produto durante a estocagem, assim ele poderá ser cultivado e vendido sem supervisão (Waltz, 2016).

Recentemente, a Autoridade Alemã concluiu que um evento de canola gerado através da tecnologia de mutagênese direcionada de oligonucleotídeo (*oligonucleotide directed mutagenesis* – ODM) (Beetham et al., 1999; Zhu et al., 1999), uma das primeiras técnicas de edição de genoma, não constitui um OGM, uma vez que é indistinguível de produtos obtidos por mutagênese convencional (BVL, 2015).

A comunidade científica e os reguladores vêm discutindo os produtos gerados por técnicas de edição de genoma amplamente (*European Food Safety Authority Panel on Genetically Modified Organisms*, 2012; *International Life Sciences Institute*, 2013; Lusser and Davies, 2013; Pauwels et al., 2014; Podevin et al., 2012, 2013), e o consenso resultante é que o grau de regulamentação destes produtos será definido pela existência ou não de integração de

sequências heterólogas via recombinação homóloga, pelas características finais do produto e pela sua utilização, bem como pela estrutura legislativa presente na região do seu lançamento (Wolt et al., 2016).

Pauwels et al. (2014) aventam que o quadro regulatório utilizado para acessar a biossegurança dos OGMs tradicionais compreendendo desde a caracterização molecular e fenotípica do evento até a detecção de efeitos indesejáveis devido à inserção do transgene em locais não controlados não seria adequado para organismos nos quais apenas *indels* ou mutações pontuais foram introduzidas por NSD1. Entretanto, Pauwels et al. (2014) e Huang et al. (2016) acreditam que a caracterização molecular da região onde ocorreu a clivagem, bem como a identificação de potenciais locais de clivagem fora do alvo, permanecem aspectos importantes que devem ser avaliados para a antecipação de efeitos adversos potenciais indesejados. Caplan et al. (2015) advertem que CRISPR tornou-se uma tecnologia popular em todo o mundo. Isto exige que existam padrões internacionais para testar a biossegurança dos organismos geneticamente editados e para possibilitar a atribuição de responsabilidade por danos.

Uma harmonização da legislação internacional, com base em uma avaliação científica, é necessária sobre esta questão, a fim de aproveitar o potencial desta tecnologia com segurança e aceitação social. O importante e desafiador é garantir que as regulamentações e as avaliações de risco exigidas permaneçam proporcionais ao nível de risco que estes organismos possam representar para a saúde humana, animal e do ambiente. Criando assim um ambiente legislativo que evite os possíveis riscos da tecnologia, mas que também estimule a inovação necessária para satisfazer outros objetivos políticos, como a segurança alimentar, o desenvolvimento econômico e a confiança do consumidor (Podevin et al., 2012).

4. Ambientais

4.1 Introdução

A *Mutagenic Chain Reaction* (MCR), ou reação mutagênica em cadeia (em tradução livre), é uma técnica de manipulação genética que permite a edição contínua de genes, enquanto estes ainda se encontram intactos em seus cromossomos de origem (para detalhes técnicos, vide capítulo 10; Gantz e Bier, 2015). Considerada um avanço científico de imenso potencial, a MCR permite que “mutações autocatalíticas” se espalhem do cromossomo de origem ao cromossomo homólogo, ocorrendo em células somáticas e da linhagem germinativa, alterando completamente o organismo em questão (Gantz e Bier, 2015).

A MCR foi inicialmente desenvolvida em 2015 por Valentino M. Gantz e Ethan Bier em experimentos com *Drosophila* (Gantz e Bier, 2015). De forma simplificada, para o funcionamento dessa técnica são necessários três componentes básicos: (i) a enzima Cas9, responsável pela clivagem do DNA-alvo, (ii) um gRNA que direciona a nuclease ao alvo e (iii) um par de sequências homólogas às sequências flanqueadoras (imediatamente adjacentes) ao local a ser editado no gene-alvo.

Dessa forma, a cassete MCR é capaz de se integrar em uma região específica do genoma (por exemplo, dentro de um gene, inativando-o) e em seguida alterando a região correspondente no cromossomo homólogo (isto é, o outro alelo) (figura 1). De acordo com o princípio básico dessa técnica, a MCR pode ser considerada uma *forma acelerada de herança*

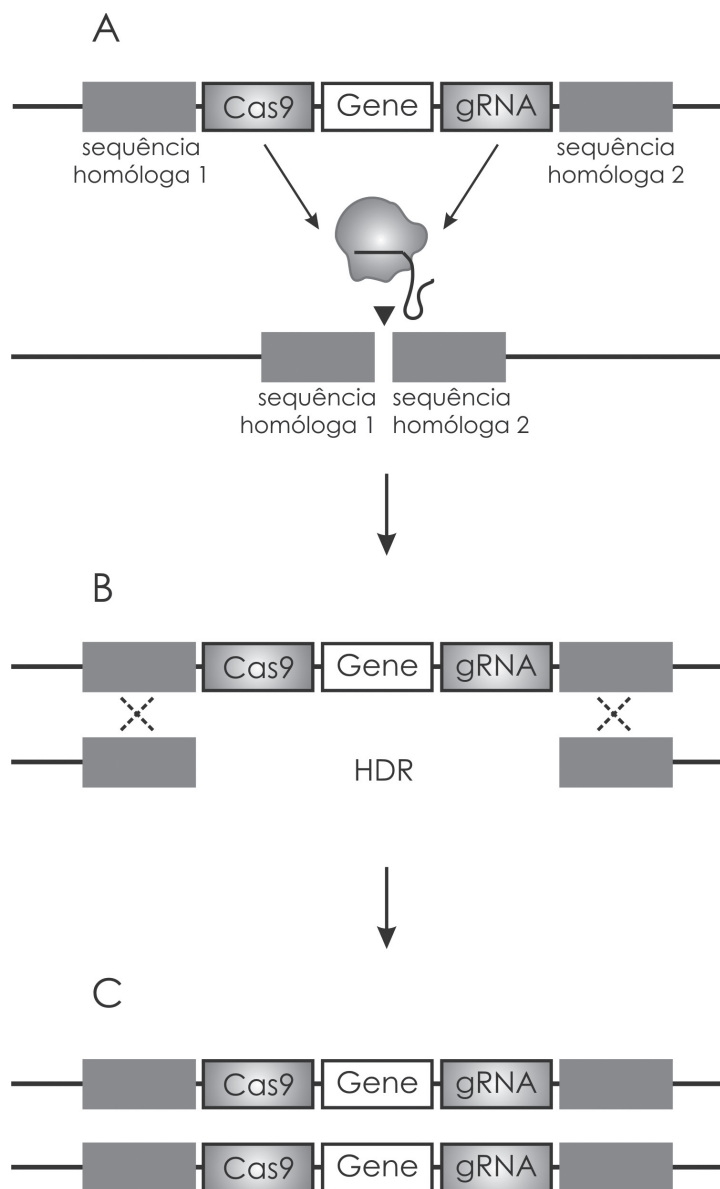


Figura 1. Modelo esquemático simplificado do mecanismo de funcionamento da *Mutagenic Chain Reaction* (MCR). Esta estratégia permite que uma “construção MCR” (genes para Cas9 e gRNA flanqueados por sequências homólogas a um loco no genoma) possa se integrar de maneira homozigótica em apenas uma geração. Eventuais (trans)genes de interesse podem fazer parte da “construção MCR” e serem herdados conjuntamente. Imagem: Pereira, TC (2016), baseado em Gantz e Bier (2015).

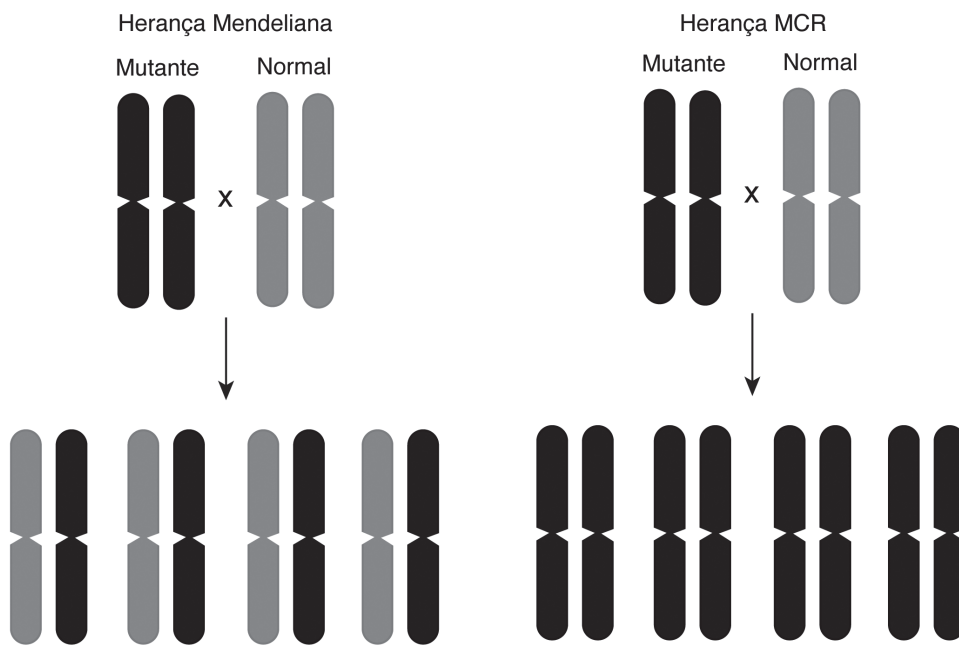


Figura 2. Modelo esquemático da transmissão de um alelo mutado através de herança mendeliana e *Mutagenic Chain Reaction*. No modelo de herança mendeliana, um organismo carregando uma cópia de um gene alterado (em preto), ao se reproduzir com um organismo não portador da alteração, normalmente passa a mutação para 50% de seus descendentes de primeira geração. O modelo de *Mutagenic Chain Reaction* (MCR) teoricamente garante que toda a descendência herde a mutação na primeira geração. Imagem: França, CT e Oliveira, MBM (2016).

mendeliana (figura 2). Isto porque, em vez de serem necessárias várias gerações para se propagar, uma mutação inicialmente em heterozigose (apenas um alelo alterado) pode entrar em homozigose em um único ciclo, potencialmente levando à manifestação do fenótipo mutante ainda na geração “0” (isto é, no organismo derivado do embrião geneticamente modificado) (Gantz e Bier, 2015).

4.2 Potenciais aplicações da técnica MCR

A tecnologia da MCR é uma poderosa ferramenta de engenharia genética, aplicável a diversos organismos e que pode ser útil para redução de problemas globais de saúde, agricultura e conservação. No momento, uma das principais utilizações da técnica é o controle de vetores de doenças infecciosas.

Gantz e colaboradores utilizaram a MCR como estratégia para introduzir um transgene no mosquito *Anopheles stephensi*, transmissor do *Plasmodium falciparum* e causador da malária (Gantz et al., 2015). A proposta da pesquisa foi gerar uma linhagem transgênica com capacidade reduzida de transmitir o patógeno. Note que a MCR é eficiente não apenas para *introduzir* um transgene no genoma, mas também para *disseminá-lo* na população: ~99,5% da prole apresentou o transgene. A expectativa é de que, no futuro, uma linhagem de mosquitos geneticamente modificada carreadora de transgenes que tornam o vetor resistente ao *P. falciparum* seja criada, auxiliando a interrupção da transmissão dessa enfermidade.

Abordagens similares também podem ser potencialmente aplicadas a pragas e espécies invasoras, bem como a introdução de mutações que conferem susceptibilidade a inseticidas, por exemplo, auxiliando na redução de problemas ambientais (Bourtzis e Hendrichs, 2014).

4.3 Cuidados no uso da técnica e regulamentações vigentes

A possibilidade de propagar transgenes ou alelos mutantes (dependendo da construção da MCR) de forma tão eficiente, mas que podem alterar populações naturais inteiras em um curto período de tempo, vem acompanhada de sérios debates sobre segurança. Apesar dos muitos benefícios, a técnica teoricamente também possui enorme potencial de causar danos até então sem precedentes em caso de liberação acidental ou intencional de organismos alterados (Esvelt et al., 2014).

Cientistas trabalhando com MCR (ou qualquer outra técnica de derivação de genes, ou seja, que permitem modificação genética ao longo do tempo) são responsáveis por, no mínimo, mantê-los confinados através de barreiras físicas, uma vez que as consequências da liberação acidental de organismos modificados no meio ambiente são imprevisíveis (Gantz e Bier, 2015; Oye et al., 2014; Esvelt et al., 2014; Burt, 2003).

O uso de barreiras físicas de contenção, em conjunto com barreiras moleculares (como métodos de reversão ou eliminação da mutação introduzida) são estratégias que devem ser utilizadas para limitar a reprodução do organismo no ambiente (Akbari et al., 2015; Esvelt et al., 2014; Henkel, Miller e Weyant, 2012). Outra alternativa são as barreiras ecológicas (como a realização de experimentação longe do *habitat* onde a população natural do organismo em estudo existe ou o uso de organismos que não possam se reproduzir na natureza). Todos esses mecanismos separados, ou em conjunto, têm sido sugeridos como forma de reduzir ainda mais os riscos de liberação acidental de organismos alterados através da MCR (Akbari et al., 2015; Caliando et al., 2015; DiCarlo et al., 2015; Oye et al., 2014).

Adicionalmente, é indicado que organismos geneticamente alterados pela MCR com potencial de propagar transgenes/alelos mutantes não sejam transferidos entre instituições, até que guias formais de biossegurança sejam estabelecidos. Em vez do transporte de amostras, os laboratórios devem, quando possível, compartilhar as construções de DNA (*DNA constructs*) ou informações e protocolos suficientes para reprodução de experimentos (Akbari et al., 2015).

Os riscos relacionados ao uso desta tecnologia pedem uma cuidadosa avaliação por parte das autoridades de biossegurança relevantes. No entanto, apesar do consenso de que todos os experimentos laboratoriais devam empregar rigorosas estratégias de confinamento, diretrizes formais (como regulamentações ou protocolos oficiais) para avaliação da eficácia de tais estratégias ainda são inexistentes ou estão em fases iniciais de desenvolvimento. Autoridades de biossegurança são, então, encorajadas a buscar orientações externas de especialistas e fazer a sua avaliação (Akbari et al., 2015; NRC, 2015; Oye et al., 2014).

4.4 Riscos ambientais da dispersão de genes derivados de MCR

Até o momento não há como prever com precisão as consequências da liberação de um organismo com uma mutação derivada de MCR para sua ou outras espécies no ambiente natural (Oye et al., 2014). No caso de dispersão de um organismo geneticamente alterado através da MCR, acredita-se que a população com o transgene/alelo mutante passaria inicialmente por um crescimento exponencial. O tempo necessário para que 100% dos indivíduos passem a carregar a mutação é, no entanto, altamente variável e dependente da

dinâmica e comportamento da população (exemplo: ciclo reprodutivo do organismo), do número de organismos mutantes introduzidos na população e da eficiência/dinâmica do fluxo de dispersão da mutação (Unckless et al., 2015a; Unckless et al., 2015b).

Apesar de ser difícil generalizar, a princípio todas as espécies que se reproduzem sexualmente, como insetos, animais e plantas podem ser alteradas por meio da MCR, e as espécies mais susceptíveis a risco de dispersão acidental são as que se reproduzem mais rapidamente (como mosquitos). Como a MCR só funciona em organismos que se reproduzem sexualmente, vírus, bactérias e a maioria dos organismos unicelulares que se reproduzem por divisão clonal ou autofertilização não podem ser alterados, e consequentemente não correm risco em caso de dispersão acidental (Esvelt et al., 2014).

Felizmente, cientistas consideram que a MCR seria ineficiente em alterar populações humanas devido ao longo período de reprodução. Em casos de contato acidental com um organismo alterado liberado na natureza, acredita-se que não muito mais do que raras reações alérgicas (devido ao contato com a proteína Cas9) poderiam ocorrer (Oye et al., 2014).

Plantações, por outro lado, seriam possivelmente susceptíveis a alterações por meio do contato com outras plantas alteradas por MCR. No entanto, a dispersão de qualquer mutação requer que esta seja eficientemente introduzida ao *pool* genético da população, e, atualmente, a maioria das plantações são cultivadas sob condições rigidamente controladas (similarmente para gado e outros animais domesticados, que são parte de programas controlados de reprodução), de modo que estas seriam rapidamente destruídas quando identificadas (Esvelt et al., 2014; Oye et al., 2014). Adicionalmente, certas alterações, como traços prejudiciais aos organismos, provavelmente necessitariam de constante reintrodução para que fossem mantidas, uma vez que estas têm custos para os indivíduos e eventualmente tendem a ser substituídas (Unckless et al., 2015a; Unckless et al., 2015b; Esvelt et al., 2014).

4.5 Controvérsias sobre a técnica MCR e sua segurança

Apesar de a CRISPR ser considerada um dos maiores avanços científicos da atualidade, há aqueles que acreditam que o estudo original que descreve a MCR (Gantz e Bier, 2015) nunca deveria ter sido publicado, uma vez que ele não inclui outras medidas de contenção além de barreiras físicas para evitar a propagação não intencional das mutações, podendo então incentivar um padrão de comportamento que é muito inferior ao recomendando (Akbari et al., 2015; Bohannon, 2015; Gantz et al., 2015). Felizmente, estudos mais recentes já fazem uso de medidas de segurança adicionais.

Por exemplo, em um estudo com *Saccharomyces cerevisiae*, a Cas9 e o gRNA foram integrados às cepas separadamente (em diferentes cassetes), o que dificulta a propagação das mutações em caso de liberação acidental. Adicionalmente, a restrição do uso da MCR exclusivamente para sequências sintéticas que só estão presentes em populações de laboratório foi outra sugestão dos autores do estudo (DiCarlo et al., 2015).

Já em cepas de *Escherichia coli*, a inclusão de um dispositivo codificado geneticamente é capaz de degradar partes predefinidas do DNA, levando as células manipuladas à morte, uma vez que o organismo esteja fora dos limites do laboratório (alterações do meio de cultivo), diminuindo assim o risco de liberação acidental (Caliando et al., 2015). Os autores do estudo original ressaltam, no entanto, que cada experimento é único, e medidas de segurança usadas em leveduras, por exemplo, podem não necessariamente ser úteis em outros organismos (Bohannon, 2015; Gantz et al., 2015).

Modelos estatísticos sugerem que, uma vez liberado, um organismo alterado pela MCR pode ser muito eficaz em alterar/substituir uma população natural, e que existem condições ótimas em que a introdução de um único indivíduo pode levar à fixação de um alelo mutante, mesmo que este possua consequências deletérias para o indivíduo. Desse modo, cientistas acreditam que a velocidade em que o processo de espalhamento da mutação pode ocorrer é motivo para precaução antes de se considerar o uso da técnica e a possível liberação de um organismo modificado no meio ambiente, seja ela acidental ou intencional, como o uso de mosquitos alterados para o combate à malária, dengue e outras doenças (Unckless et al., 2015a; Unckless et al., 2015b). Adicionalmente, mesmo que as alterações genéticas possam ser revertidas, qualquer população natural consequentemente alterada teria sua diversidade genética reduzida, podendo se tornar mais vulnerável a pressões naturais ou antropogênicas (Oye et al., 2014).

4.6 Conclusão

Os riscos e consequências ambientais da dispersão de transgenes/alelos mutantes via MCR dependem de diversos fatores, como a espécie envolvida e a ocorrência de novas mutações na mutação inicial feita via MCR, entre outras (Esvelt et al., 2014; Oye et al., 2014). Os potenciais usos da MCR são únicos, e é importante que cada um seja ponderado e discutido com transparência e cautela. Claramente, mais pesquisas são necessárias para que possamos compreender como indivíduos, populações e ecossistemas irão responder a alterações tão dramáticas.

5. Pioneirismo

5.1 As patentes e o mercado da biotecnologia

Atualmente, um tema vem recebendo notoriedade no meio acadêmico, devido aos desafios presentes na comunidade científica, e refere-se à importância do pioneirismo na pesquisa. Segundo a definição encontrada nos dicionários, o “Pioneiro” é quem primeiro desbrava regiões incultas (desbravador). Esta descrição sintetiza a rotina diária de um pesquisador dedicado à busca por inovações e novas criações tecnológicas que possam vir a ser patenteadas. Primeiramente, em face da relevância do assunto, a presente seção tem como objetivo apresentar uma breve elucidação sobre patentes, tendo como única base a área de biológicas com ênfase no sistema CRISPR-Cas9, assunto central deste livro.

Historicamente é inegável a relevância do potencial das descobertas biotecnológicas e o seu impacto global. Ao longo dos últimos anos não faltam exemplos que mostram os notáveis progressos. Em 1953, James Watson e Francis Crick apresentam a estrutura do DNA (Watson e Crick, 1953). Em 1985 surge a PCR (*Polymerase Chain Reaction*), e alguns anos depois dá-se o início do projeto Genoma Humano. O ano de 1996 é marcado pelo nascimento da ovelha Dolly, primeiro mamífero a ser clonado. Dois anos mais tarde, Andrew Fire e colaboradores desenvolvem a técnica de RNAi (Fire et al., 1998). No Brasil, no ano 2000, anuncia-se o sequenciamento do genoma da bactéria *Xylella fastidiosa*, e em 2012 foi descrita pela primeira vez a tecnologia de edição gênica CRISPR-Cas9 por Jennifer Doudna e Emmanuelle Charpentier (Jinek et al., 2012).

A lista de descobertas é extremamente extensa, no entanto, há ainda inúmeras áreas a serem exploradas. Grupos de pesquisas de todo o mundo têm estudado excessivamente mecanismos biológicos vislumbrando possíveis aplicabilidades ainda não descritas passíveis de uma publicação científica ou proteção através de um documento de patente.

A pesquisa voltada para as moléculas com atividades biológicas e/ou farmacológicas, propriedades físico-químicas otimizadas (melhores do que as existentes) e biomateriais, alvos centrais da indústria biotecnológica, pode tornar-se parte de um emergente mercado global que cresceu no período de 2002-2006 a uma taxa anual (CAGR – *Compound Annual Growth Rate*) de 13,4%, partindo de uma base de US\$ 92,9 bilhões em 2002 a US\$ 153,5 bilhões em 2006. Ainda em 2006, o setor de medicina e saúde humana foi o que mais se destacou, gerando receitas de US\$ 96,2 bilhões, equivalentes a 62,5% do valor do mercado global de biotecnologia (Datamonitor, 2007; Global Bioeconomy Consulting LLC, 2007). Já em 2016, estes mesmos dados revelam uma receita que atinge US\$ 1 trilhão, com crescimento do setor de 3% ao ano (2011-16). Somente nos Estados Unidos, considerado líder no segmento, mais de 810 mil profissionais trabalham na indústria biofarmacêutica, que gera um total de 3,4 milhões de empregos indiretos (SelectUSA, 2016).

Diante de todos esses indicadores positivos, indústria e pesquisa científica foram estimuladas pelo cenário promissor, firmando assim algumas parcerias de desenvolvimento e produção em grande escala de itens de origem biológica.

Em decorrência desta progressiva expansão do mercado biotecnológico, houve um considerável aumento dos depósitos de patentes em relação a estas tecnologias, visando principalmente garantir a proteção sobre o monopólio daquele produto e frequentemente projeção internacional.

Dentro deste cenário, fulgura a recém-descoberta técnica de CRISPR-Cas9, ferramenta que, conforme já retratado, abre inúmeras possibilidades de aplicações e inovações biotecnocientíficas, tornando-o altamente valioso em nível mercadológico e, portanto, alvo de uma série de conflitos não usuais que serão descritos a seguir.

5.2 A disputa Doudna-Charpentier versus Zhang

Relembrando alguns pontos-chave fundamentais para a compreensão da tratativa, o potencial de CRISPR-Cas9 como uma ferramenta de edição genética foi descrito pela primeira vez em uma publicação de agosto de 2012 na revista *Science* (Jinek et al., 2012), trabalho este liderado por Jennifer Doudna, da Universidade da Califórnia em Berkeley, e Emanuelle Charpentier, do Centro de Pesquisa de Infecções Helmholtz, Alemanha (Sherkow et al., 2015). As duas pesquisadoras submeteram o depósito de patente para o sistema CRISPR-Cas9 em maio de 2012, que incluía cerca de 155 reivindicações de patentes consistindo em diversos tipos de aplicabilidades da nova ferramenta em diferentes tipos celulares (Sherkow et al., 2015). Este pedido encontra-se sob análise do USPTO (*United States Patent e Trademark Office*), escritório americano responsável pela análise dos depósitos de patentes sob a numeração PCT/US2013/032589 (2013).

Em paralelo, em dezembro do mesmo ano, o pesquisador Feng Zhang, do MIT (*Massachusetts Institute of Technology*), também submeteu o pedido de patente referente à utilização do sistema CRISPR-Cas9 em células eucarióticas (No. 2014/0336133 (2014)), sendo este já concedido em abril de 2014. Além desses dois pedidos de patentes, há uma série de outras solicitações envolvendo o sistema CRISPR-Cas9.

No entanto, há um impasse de grandes proporções envolvendo Doudna-Charpentier e Zhang referente à disputa pela aquisição de patente, gerando desafios que acometem não só os pesquisadores em questão, bem como toda a comunidade da área biotecnológica.

O primeiro ponto a ser delineado, que levantou uma série de discussões pertinentes abrangendo algumas jurisdições de diferentes países, relaciona-se à questão do pioneirismo. Este ponto chama a atenção nesse caso específico, pois, levando em consideração as datas de submissão para a aquisição de patente, Doudna-Charpentier fizeram o pedido antes de Zhang, e este último obteve a concessão, enquanto as pesquisadoras ainda possuem seu pedido sob análise (Ledford, 2016).

O pedido de solicitação de patente de ambos os pesquisadores foi depositado nos Estados Unidos, líder mundial nas descobertas científicas e inovações, tradicionalmente um dos maiores mercados para a biotecnologia. De acordo com a Lei de Proteção à Propriedade Intelectual americana, os detalhes de como uma patente deve ser examinada estão no Manual de Procedimentos de Exame de Patente – *Manual of Patent Examining Procedure* (USPTO, 2006a). Brevemente será descrito nas próximas linhas a atual regra que regula o sistema de patente dos EUA inerente ao tema em pauta.

Segundo os critérios de patenteabilidade americanos, as patentes podem ser obtidas por quem quer que invente ou descubra qualquer processo ou composição da matéria que seja nova (35 USC 102), útil e não óbvia (35 USC 101) (Webber et al., 2014). Considerado um dos sistemas menos complexos de depósitos de patentes, quando comparado com o modelo de outros países, o sistema americano modificou suas regras em março de 2013. Antes desta data, vigorava a regra de que a patente seria concedida para o inventor pioneiro, porém, foi modificada, e atualmente quem obtém o privilégio da patente é quem primeiro realiza o depósito de solicitação de patente.

De um lado, Doudna-Charpentier realizam o depósito da patente em maio de 2012, que antecede a data de mudança da regra; de outro lado, Zhang deposita também no mesmo ano, ou seja, ambos sob a mesma legislação. Dessa forma, Zhang alega ter sido o primeiro inventor e Doudna-Charpentier afirmam que, além de serem as primeiras inventoras, também se enquadram na nova lei em vigor, sendo as primeiras depositantes. Todavia, o processo ocorre em andamento sob o regime do antigo modelo fundamentado no pioneirismo.

Além da rivalidade no âmbito acadêmico, ambos os pesquisadores estão relacionados com a fundação de empresas *Start-ups* voltadas para a edição gênica. *Editas Medicine* ganhou a primeira patente (n° 8697359 - Sistema CRISPR-Cas e métodos de alteração de expressão de produtos gênicos) neste campo utilizando o sistema CRISPR-Cas, tendo como cofundador Feng Zhang. A empresa em questão foi fundada pela união de duas instituições parceiras: *Broad Institute* e MIT, sendo as duas de Cambridge, Massachusetts. Em 2013, *Editas* foi lançada com US\$ 43 milhões em financiamentos de empresas como *Flagship Ventures*, *Polaris Partners* e *Third Rock Ventures*, incluindo a participação da Fundação de Parceiros da Inovação (Sheridan, 2014).

A *Start-up CRISPR Therapeutics*, sob a direção científica de Doudna-Charpentier, também se encontra na disputa do concorrido mercado de edição gênica. Apesar de Zhang ter conseguido a patente inicial, este não se encontra totalmente imune, pois o depósito das pesquisadoras ainda está sob análise na USPTO (PCT/US2013/032589) e pode ser aceito. Ainda assim, caso não o seja, as duas podem entrar com um recurso denominado “*interference proceeding*” (35 USC 135) (Patentek, 2016) contra a patente inicial de Zhang.

Assim, existem vários desdobramentos que podem ser previstos: se Doudna-Charpentier ganharem o pedido de interferência, a patente de Zhang seria anulada, embora não haja garantia de que o depósito delas ainda em revisão seja aceito. Em contrapartida, se os argumentos de Zhang se sobrepuserem ao das pesquisadoras, ele mantém a sua patente e elas provavelmente perderiam a sua solicitação (Sherkow et al., 2015).

Ainda para fomentar a discussão envolvendo a disputa da patente principal de CRISPR-Cas9, se a patente de Zhang for invalidada, todas as patentes secundárias a esta podem tornar-se irregulares, inclusive produtos que já tenham sido regulamentados pelo FDA (*Food and Drug Administration*) (Sherkow et al., 2015). Outro aspecto que merece certa relevância é que, depois que um mecanismo torna-se amplamente compreendido e conhecido, como a PCR, ele pode converter-se em óbvio e, portanto, não ser elegível a patente e ser invalidado (Sherkow et al., 2015; Ledford, 2016).

Em abril de 2015, apesar das possíveis consequências negativas conforme citado, o grupo de Berkeley decidiu entrar com o pedido de interferência sob a patente de Zhang na USPTO, este responsável por decidir quem dos pesquisadores rivais seria o primeiro inventor da técnica de edição gênica. Dessa forma, o escritório de patentes iniciou o processo de interferência de patente, agindo de modo equivalente a um processo judicial, solicitando evidências, depoimentos comprovantes de ambos os lados, publicações e cadernos de laboratório para embasamento da decisão (Ledford, 2016; Nature Biotechnology News, 2016).

Como se já não houvesse tantas considerações a serem delineadas e entraves a serem solucionados em relação à situação da patente CRISPR-Cas9 nos Estados Unidos, ainda há a questão europeia a ser definida e a dúvida proeminente de quem seria então o proprietário intelectual na Europa.

Os direitos de propriedade industrial são de efeito local. Para ter sua proteção estendida em outros territórios (extensão internacional), é necessário realizar a solicitação de um pedido de patente na instituição provedora da outra região.

Adentrando as regras de outro continente, a Europa é representada pelo Escritório Europeu de Patentes (EPO), que tem como objetivo estabelecer as regras de proteção patentária para todos os 31 países constituintes (INPI, 2007), e é o quinto maior escritório de patentes do mundo, segundo a WIPO (*World Intellectual Property Organization*) (WIPO, 2016). A proteção das invenções biotecnológicas é citada nos artigos 52 e 53 da Convenção de Patentes Europeia e pela Diretiva Europeia 98/44EC.

Assim, atendendo aos critérios estabelecidos pelas supracitadas, o pesquisador Zhang do MIT prossegue com a aplicação da patente conquistada na USPTO na Europa. No entanto, o depósito não chegou a ser analisado, inclusive não sendo encaminhado para a Divisão de Exame. Pois, após a EPO receber a solicitação de patente, houve a intervenção de uma observação de terceiros (*third-party*) similar a uma denúncia realizada de forma anônima que paralisou o andamento do processo (Kupecz, 2014).

O artigo 115 da EPC (*European Patent Convention*) prevê que antes da solicitação do pedido de patente é possível que qualquer terceiro apresente argumentos que inviabilizem a continuidade do processo, provocando atrasos e podendo até mesmo culminar na recusa da concessão de uma patente (Kupecz, 2014). Neste caso em referência, o relatório de terceiros apresentado à EPO contemplava cerca de 20 incongruências técnicas em oposição à patente de Zhang, sendo que algumas destas já haviam sido previamente investigadas pela própria EPO. Um dos argumentos descritos neste documento mencionava a ausência de invenção e novidade (Kupecz, 2014).

Conforme citado anteriormente, o regime de aquisição de patentes adotado pelos Estados Unidos é simplista quando comparado às regras de outros países, e isso inclui o sistema europeu, dotado de exigências e regras mais complexas.

De acordo com as regras existentes e praticadas pela EPO, a data de depósito de um pedido de patente dá origem ao direito de prioridade. Assim, o requerente dispõe de um prazo de 12 meses a contar desta data para apresentá-lo em outros países, mantendo como data de depósito dos pedidos posteriores a do primeiro depósito (data de prioridade).

Desta forma, se os argumentos do relatório de terceiros (*third-party*) forem aceitos, pode ocorrer uma alteração no que concerne às datas de depósito de Zhang (MIT) (de 12 de dezembro de 2012 para 12 de dezembro de 2013), o que culminará na alteração dos documentos solicitados para o depósito. Além disso, o depósito da patente deve ser feito anteriormente à publicação; se ocorrer a mudança de datas, haverá uma inversão e a publicação se tornará anterior ao pedido do depósito, podendo assim ser interpretado pelo comitê europeu como ausência de novidade.

Diante de todos os fatos, o processo relatado encontra-se única e exclusivamente sob a análise da Divisão de Exames da EPO, cabendo a ela a decisão da concessão da patente CRISPR-Cas9 para o pesquisador Zhang. Durante o decorrer do processo, é possível que novas evidências sejam incorporadas tanto pelo requerente quanto por terceiros ou pela própria EPO.

Como consequência, dificilmente é possível fazer uma previsão em relação ao futuro deste processo. Caso o pedido de aquisição de patente seja negado pela Divisão de Exame, existe a possibilidade da abertura de um pedido de recurso pelo depositante no Conselho de Recursos da EPO, um processo com a duração mínima de dois anos. Ainda assim, se através do recurso a patente for concedida, o processo continua em aberto por mais nove meses, ficando disponível para a intervenção de terceiros. Neste ponto presume-se que haverá oposição sobre a patente de Zhang, prorrogando o processo por aproximadamente mais cinco anos. Portanto, a tecnologia CRISPR-Cas9 ainda não se encontra patenteada na Europa e o seu processo ainda demorará alguns anos até que possa finalmente ser concluído.

Apesar de toda a incerteza que norteia o futuro de uma das tecnologias mais revolucionárias dos últimos tempos e forte candidata ao Nobel, em nível mercadológico CRISPR-Cas9 é absoluto sinônimo de investimentos. A gigante *Bayer* investiu neste segmento, por meio de uma *joint-venture* contratual formada com a *CRISPR Therapeutics* através de um consórcio de empresas. Este tipo de associação tem como principal característica o benefício mútuo das companhias envolvidas, como o compartilhamento do conhecimento e novas tecnologias, competitividade internacional e ampliação da rede estratégica de negócios. A *Bayer* está investindo mais de US\$300 milhões em Pesquisa e Desenvolvimento (P&D) para os próximos cinco anos neste segmento. A *DuPont*, segunda maior empresa química do mundo com forte liderança em ciência e tecnologia, também formou uma aliança com a *Caribou Biosciences* (laboratório da Universidade da Califórnia, liderado por Jennifer Doudna) envolvendo aplicações da ferramenta CRISPR-Cas9 em plantas, resultando em publicações (Grushkin, 2016).

Em 2014, Jennifer Doudna e Emmanuelle Charpentier foram indicadas ao Prêmio Nobel de Química, como o esperado. No entanto, os laureados nesta premiação foram os pesquisadores Eric Betzig, Stefan W. Hell e William E. Moerner. Mas, visto que mais da metade dos laureados esperaram cerca de 20 anos para ganhar, as pesquisadoras continuam na disputa (Cesare, 2015).

6. Conclusões

A constatação de que um achado científico é de “alto impacto” pode ser feita de formas além da compilação do número de artigos e citações a ele associados. Pode-se dizer que a descoberta alcançou um nível extraordinário de relevância quando ela suscita importantes debates dentro da comunidade não acadêmica, desperta um intenso interesse do setor privado e força os legisladores a reverem conceitos.

Referências

- Akbari OS, Bellen HJ, Bier E, Bullock SL, Burt A, Church GM, et al. BIOSAFETY. Safeguarding gene drive experiments in the laboratory. *Science*. 2015;349(6251):927-9.
- Araki M, Ishii T. Towards social acceptance of plant breeding by genome editing. *Trends Plant Sci*. 2015;20(3):145-9.
- Araki M, Nojima K, Ishii T. Caution required for handling genome editing technology. *Trends Biotechnol*. 2014;32:234-7.
- Aristóteles. *Ética a Nicômaco*. 5. ed. São Paulo: Martin Claret; 2011. Coleção a Obra Prima de Cada Autor. Tradução: Torrieri Guimarães.
- Beetham PR, Kipp PB, Sawycky XL, Arntzen CJ, May GD. A tool for functional plant genomics: Chimeric RNA/DNA oligonucleotides cause in vivo gene-specific mutations. *Proc Natl Acad Sci USA*. 1999;96:8774-8.
- Bohannon J. Biologists devise invasion plan for mutations. *Science*. 2015;347(6228):1300.
- Bortesi L, Fischer L. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol Adv*. 2015;33:41-52.
- Bourtzis K, Hendrichs J. Development and evaluation of improved strains of insect pests for sterile insect technique (SIT) applications. *BMC Genet*. 2014;15(Suppl 2):11.
- Burachik M. Regulatory framework for biotechnology-derived crops with specific focus on new plant breeding techniques in Argentina. Comparative Regulatory Approaches for New Plant Breeding Techniques, European Commission workshop; 2011 sept 12-13; Seville, Spain.
- Burt A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc Biol Sci*. 2003;270(1518):921-8.
- BVL. The website of the Federal Office of Consumer Protection and Food Safety (BVL) of the German government. 2015. Available from: http://www.bvl.bund.de/DE/06_Gentechnik/04_Fachmeldungen/2015/2015_06_03_Fa_CIBUS.html.
- Caliando BJ, Voigt CA. Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nat Commun*. 2015;6:6989.
- Callaway E. Second Chinese team reports gene editing in human embryos. Study used CRISPR technology to introduce HIV-resistance mutation into embryos. *Nature*, US, 08 April 2016. News and comments, Article [cited 2016 Junho 20]. Available from: <http://www.nature.com/news/second-chinese-team-reports-gene-editing-in-human-embryos-1.17118>
- Camacho A, Van Deynze A, Chi-Ham C, Bennett AB. Genetically engineered crops that fly under the US regulatory radar. *Nat Biotechnol*. 2014;32:1087-91.
- Caplan AL, Parent B, Shen M, Plunkett C. No time to waste: the ethical challenges created by CRISPR. *EMBO Rep*. 2015;16(11):1421-6.
- Cesare C. Nobel speculation kicks into high gear. *Nature*. 2015.
- Cruz MR, Oliveira SLT, Portillo JAC. A Declaração Universal sobre Bioética e Direitos Humanos: contribuições ao Estado brasileiro. *Revista Bioética*. 2010;18(1):93-107.
- Curtin SJ, Voytas DF, Stupar RM. Genome engineering of crops with designer nucleases. *Plant Genome*. 2012;5:42-50.
- Cyranoski D. Scientists sound alarm over DNA editing of human embryos experts call for halt in research to work out safety and ethics issues. *Nature*, US, 12 March 2015, News and comments, Article [cited 2016 Junho 24]. Available from: <http://www.nature.com/news/scientists-sound-alarm-over-dna-editing-of-human-embryos-1.17110>
- Cyranoski D, Reardon S. Embryo editing sparks epic debate In wake of paper describing genetic modification of human embryos, scientists disagree about ethics. *Nature*. 2015;520:593-4.
- DATAMONITOR. Global Biotechnology: industry profile. 2007.
- DiCarlo JE, Chavez A, Dietz SL, Esvelt KM, Church GM. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol*. 2015;33(12):1250-5.
- Esvelt KM, Smidler AL, Catteruccia F, Church GM. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife*. 2014:e03401.
- European Food Safety Authority Panel on Genetically Modified Organisms – EFSA. Scientific opinion addressing the safety assessment of plants developed using zinc finger nuclease 3 and other site-directed nucleases with similar function. *EFSA J*. 2012;10:2943.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998;391:806-11.
- Gantz VM, Bier E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*. 2015;348(6233):442-4.
- Gantz VM, Jasinskiene N, Tataronkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc Natl Acad Sci USA*. 2015;112(49):E6736-43.

- Global Bioeconomy Consulting LLC. Global Hubs and Global Nodes of Biotechnology: an international scan of biotechnology strategies, initiatives and institutional capacity. Washington; 2007.
- Goldim JR. Genetics and ethics: a possible and necessary dialogue. *J Community Genet*. 2015;6(3):193-6.
- Grushkin D. DuPont in CRISPR-Cas patent land grab. *Nat Biotechnol*. 2016;34:13.
- Hartung F, Schiemann J. Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. *Plant J*. 2014;78:742-52.
- Henkel RD, Miller T, Weyant RS. Monitoring select agent theft, loss and release reports in the United States—2004–2010. *Applied Biosafety*. 2012;17(4):171-80.
- Huang S, Weigel D, Beachy RN, Li J. A proposed regulatory framework for genome-edited crops. *Nat Genet*. 2016;48(2):109-11.
- Instituto Nacional da Propriedade Industrial – INPI. Estudo Comparativo dos Critérios de Patenteabilidade para Invenções Biotecnológicas em Diferentes Países. 2007.
- International Bioethics Committee – IBC. Report of the IBC on Updating Its Reflection on the Human Genome and Human Rights [cited 2016 Maio 10]. Available from: <http://unesdoc.unesco.org/images/0023/002332/233258E.pdf>
- International Life Sciences Institute. ILSI Workshop on New Breeding Technologies (NBT). Canberra: International Life Sciences Institute SEA Region Australasia, ISAAA; 2013 [cited 2016 Junho 24]. Available from: <http://www.isaaa.org/resources/publications/briefs/51/executivesummary/default.asp>
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna J, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816-21.
- Jones HD. Regulatory uncertainty over genome editing. *Nature Plants*. 2015;1:1-3.
- Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen T, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529:490-5.
- Kupez A. Who owns CRISPR-Cas9 in Europe? *Nat Biotechnol*. 2014;32:1194-6.
- Ledford H. Bitter fight over CRISPR patent heats up. *Nature*. 2016;529.
- Liang P, et al. CRISPR/Cas9-mediated gene editing in human triploid zygotes. *Protein Cell*. 2015;6:363.
- Lusser M, Davies HV. Comparative regulatory approaches for groups of new plant breeding techniques. *N Biotechnol*. 2013;30:437-46.
- Mahfouz MM, Piatek A, Stewart Jr CN. Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnol J*. 2014;12:1006-1-14.
- National Research Council. National Academy of Sciences. Washington: Gene Drive Research in Non- Human Organisms: Recommendations for Responsible Conduct [cited 2015 Março 11]. Available from: <http://bit.ly/CurrProjects-regul>
- Nature Biotechnology News. CRISPR patents to go on trial. *Nat Biotechnol*. 2016 Fevereiro;34:121.
- Nietzsche F. Assim falou Zarathustra. São Paulo: Companhia das Letras; 1998.
- Oye KA, Esvelt K, Appleton E, Catteruccia F, Church G, Kuiken T, et al. Biotechnology: regulating gene drives. *Science*. 2014;345(6197):626-8.
- Patentek. 5 U.S.C. 135 Interferences. – Patent Laws. [cited 2016 Abril 3]. Available from: <http://www.patentek.com/patent-interference-law-5-usc-135/>
- Pauwels K, Podevin N, Breyer D, Carroll D, Herman P. Engineering nucleases for gene targeting: safety and regulatory considerations. *N Biotechnol*. 2014;31:18-27.
- Pessini L. As origens da bioética: do credo bioético de Potter ao imperativo bioético de Fritz Jahr. *Revista Bioética*. 2013;21(1):9-19.
- Podevin N, Davies HV, Hartung F, Nogue F, Casacuberta JM. Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. *Trends Biotechnol*. 2013;31:375-83.
- Podevin N, Devos Y, Davies HV, Nielsen KM. Transgenic or not? No simple answer! *EMBO Rep*. 2012;13:1057-61.
- Potter VR. Bioethics, science of survival. *Perspect Biol Med*. 1970;14:27-153.
- Potter VR. Bioethics: bridge to the future. Englewood Cliffs: Prentice-Hall; 1971.
- Prasad KV, Song B, Olson-Manning C, Anderson JT, Lee GR, Schranz ME, et al. A gain-of-function polymorphism controlling complex traits and fitness in nature. *Science*. 2012;337:1081-4.
- Select USA. The Pharmaceutical and Biotech Industries in the United States. The International Trade Administration (ITA). [cited 2016 Junho 18]. Available from: <https://www.selectusa.gov/pharmaceutical-and-biotech-industries-united-states>
- Shan Q, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol*. 2013;8:686-8.
- Sheridan C. First CRISPR-Cas patent opens race to stake out intellectual property. *Nat Biotechnol*. 2014;32(7).
- Sherkow JS. Law, history and lessons in the CRISPR patent conflict. *Nat Biotechnol*. 2015;33:256-7.
- Siqueira JE, Diniz N. Ética e responsabilidade em genética. In: Garrafa V, Pessini L. Bioética: poder e injustiça. São Paulo: Centro Universitário São Camilo; 2003. p. 225-34.
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science*. 2016;351(6268):84-8.
- Smyth S, McHughen A. Regulating innovative crop technologies in Canada: the case of regulating genetically modified crops. *Plant Biotechnol J*. 2008;6:213-25.
- Unckless RL, Messer PW, Connallon T, Clark AG. Modeling the manipulation of natural populations by the mutagenic chain reaction. *Genetics*. 2015a;201(2):425-31.
- Unckless R, Messer P, Clark A. Theoretical consequences of the Mutagenic Chain Reaction for manipulating natural populations. *bioRxiv*. 2015b;1:018986.
- UNESCO. UNESCO panel of experts calls for ban on “editing” of human DNA to avoid unethical tampering with hereditary traits. [cited 2016 Junho 24]. Available from: http://www.unesco.org/new/en/media-services/single-view/news/unesco_panel_of_experts_calls_for_ban_on_editing_of_human_dna_to_avoid_unethical_tampering_with_hereditary_traits/#.VzLttQrJH0

- United Nations. International Declaration on Human Genetic Data. Paris; 2003 [cited 2016 Junho 24]. Available from: <http://www.unesco.org/new/en/social-and-human-sciences/themes/bioethics/human-genetic-data/>
- United Nations. Universal Declaration on the Human Genome and Human Rights. Paris; 1997 [cited 2016 Junho 24]. Available from: <http://www.unesco.org/new/en/social-and-human-sciences/themes/bioethics/human-genome-and-human-rights>
- United Nations. The Universal Declaration of Human Rights. Paris; 1948 [cited 2016 Junho 24]. Available from: <http://www.un.org/en/documents/udhr/>
- Voytas DF, Gao C. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* 2014;12:e1001877.
- Waltz E. Gene-edited CRISPR mushroom escapes US regulation. *Nature*. <http://dx.doi.org/10.1038/nature.2016.19754>.
- Waltz E. Tiptoeing around transgenics. *Nat Biotechnol.* 2012;30:215-7.
- Watson J, Crick F. Molecular structure of nucleic acids: a structure for deoxyribose nucleic Acid. *Nature*. 1953;17:737-8.
- Webber P. Does CRISPR-Cas open new possibilities for patents or present a moral maze? *Nat Biotechnol.* 2014;32:331-3.
- World Intellectual Property Organization – WIPO. World Patent Report: a statistical review. 2008 [cited 2016 Março 31]. Available from: http://www.wipo.int/ipstats/en/statistics/patents/wipo_pub_931.html
- Wolt JD, Wang K, Yang B. The regulatory status of genome-edited crops. *Plant Biotechnol J.* 2016;14:510-8.
- Zhu T, Peterson DJ, Tagliani L, St. Clair G, Baszczyński CL, Bowen B. Targeted manipulation of maize genes in vivo using chimeric RNA/DNA oligonucleotides. *Proc Natl Acad Sci USA.* 1999;96:8768-73.

Tópicos de fronteira

Capítulo 12

Luciana Rodrigues Carvalho Barros^{1,a}, Thais Campos de Oliveira^{2,b},
Augusto Schrank^{2,b} e Tiago Campos Pereira^{3,4,c}

¹*Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), RJ - Brasil*

²*Centro de Biotecnologia, UFRGS, RS - Brasil*

³*Dept.º de Biologia, FFCLRP - USP, SP - Brasil*

⁴*Programa de Pós-Graduação em Genética, FMRP - USP, SP - Brasil*

^a*Autoria da seção: 2*

^b*Autoria da seção: 3*

^c*Autoria das seções: 4 e 5*

Estruturação do capítulo

1. Introdução
2. CRISPR *goes retro*
3. MIMIVIRE: CRISPR de vírus gigantes?
 - 3.1 Mimivírus
 - 3.2 O mecanismo MIMIVIRE
4. Tópicos especiais sobre CRISPR
5. Conclusões

1. Introdução

Apresentamos ao longo do livro várias das principais aplicações e variações da técnica de CRISPR. Neste capítulo final, abordaremos alguns temas situados nos limites desta área de conhecimento, artigos que projetam a CRISPR para outros domínios além das arqueas e bactérias, além das aplicações biotecnológicas mais convencionais e intuitivas, cujas implicações podem dar novos rumos à genética molecular.

2. CRISPR *goes retro*

Os vírus de bactérias, ou bacteriófagos, são a mais abundante forma de vida. Já foram descritos diversos mecanismos de defesa de bactérias contra essas infecções, incluindo a diminuição de receptores de entrada viral e mecanismos de indução de morte celular. Nos últimos anos tem sido descrito um sistema de imunidade adaptativa denominado CRISPR-Cas, que é o primeiro e único de tal tipo em Arqueas e Bactérias. Por meio desse sistema, procariotos têm memória dos encontros com os organismos invasores via a aquisição de pequenos fragmentos de seus elementos genéticos.

Como visto em capítulos anteriores, o sistema pode ser dividido em três etapas (vide cap. 1): a primeira consiste no reconhecimento, clivagem e incorporação de fragmentos do material genético exógeno, como DNA viral, nos locos CRISPR (adaptação). Na segunda etapa, o loco CRISPR é transcrito junto com os genes *Cas*, produzindo os crRNAs. Na última etapa, os crRNAs associados às proteínas *Cas* reconhecem elementos genéticos invasores por complementaridade da sequência de bases e clivam esses elementos (interferência), impedindo sua ação. Tanto DNA quanto RNA podem ser complementares ao crRNA e clivados pelas *Cas*. Como mecanismo de defesa, essa característica é importante para impedir a ação de vírus (com genoma em RNA ou DNA) e plasmídeos (Barrangou, 2013).

A capacidade adaptativa das bactérias depende da habilidade do loco CRISPR em incorporar novos espaçadores (Sontheimer e Marrafini, 2016). Os mecanismos de aquisição de novos espaçadores através de DNA têm sido cada vez mais profundamente elucidados. Até o momento, esses sistemas foram divididos em três tipos principais baseados nas sequências dos genes *Cas*, nas sequências repetitivas dentro do CRISPR e na organização dos *operons* das *Cas* (Marakova et al., 2015). Desses sistemas, o mais bem descrito é o Tipo II, devido a sua simplicidade, já que apenas a *Cas9* é requerida em todas as etapas e a sua grande aplicação para edição de genomas.

O reconhecimento apenas de DNA exógeno pelas proteínas *Cas* é um importante fator de prevenção de autoimunidade. Essa distinção entre o próprio (*self*) e o não-próprio (*non-self*) se dá por alguns mecanismos que ainda estão sendo estudados. O primeiro é o requerimento de uma sequência PAM presente no DNA exógeno, mas não dentro do loco CRISPR. O PAM é composto por uma sequência de três a oito bases específicas para cada nuclease, sendo “NGG” para a *Cas9* de *Streptococcus pyogenes*, por exemplo.

Outro mecanismo de prevenção de autoimunidade e clivagem dos próprios cromossomos se dá pela presença de regiões chamadas *Chi* (*crossover hotspot instigator*), que são octâmeros altamente enriquecidas no genoma bacteriano, sugerindo que limitem a aquisição de espaçadores do DNA próprio (Karoui et al., 1999; Levy et al., 2015). Outro fator é a preferência das *Cas* em clivar regiões em processo de duplicação nas regiões das forquilhas de replicação. Esse modelo explica a aquisição de espaçadores de plasmídeos de

grande número de cópias e fagos (Levy et al., 2015). Entretanto, o sistema Tipo III não requer PAM e é baseado na proximidade entre a sequência repetitiva e a sequência espaçadora do loco CRISPR (Marrafini e Sontheimer, 2010).

Até o final de 2015, acreditava-se que apenas moléculas de DNA seriam reconhecidas e incorporadas nos locos CRISPR como novos espaçadores (revisado por Marrafini, 2015). Porém, vários bacteriófagos (ou simplesmente fagos) possuem genoma na forma de RNA e alguns experimentos sugeriam que as bactérias também adquiriam resistência a esses fagos. Através do sequenciamento de locos CRISPR, foram observados vários fragmentos pequenos desses genomas (Marakova et al., 2006). Acreditava-se que o mecanismo de inserção de fragmentos de RNA aconteceria após atividade da enzima transcriptase reversa do próprio fago durante a infecção. O reconhecimento e a clivagem de moléculas de RNA foram observados no sistema do Tipo III. Porém, acreditava-se que eram apenas aqueles transcritos do DNA alvo original (Peng et al., 2014; Samai et al., 2015). Todavia, esse modelo não contemplava todas as observações, e uma questão sobre a aquisição de espaçadores tanto provenientes de RNA quanto de DNA foi levantada.

Recentemente, Silas e colaboradores (Silas et al., 2016) responderam essa questão procurando, através de ferramentas de bioinformática, enzimas com atividade de *transcrição reversa* em bactérias com sistema CRISPR Tipo III. Tais transcriptases reversas (RTs) utilizam RNA como substrato para a síntese de cDNA e são normalmente associadas a retrovírus e retroelementos. Os autores encontraram uma classe de regiões codificantes associadas ao CRISPR, na qual a *cas1* é fusionada a uma RT (RT:Cas1).

A Cas1 associada a Cas2 catalisa a integração de novos espaçadores no loco CRISPR. Utilizando sequências contendo íntrons que sabidamente sofreriam *self-splicing*, os autores mostraram a aquisição das sequências maduras no loco CRISPR. Como *splicing* só ocorre ao nível de RNA, esse experimento sugeriu que o RNA foi a fonte da aquisição do novo espaçador. Para confirmar, purificaram as proteínas RT:Cas1 e Cas2 e incubaram com RNA ou DNA sintéticos *in vitro*. Tanto DNA quanto RNA estavam ligados ao complexo RT:Cas1 e Cas2, sendo ambos incorporados. A síntese de DNA a partir do RNA fornecido também foi observada. Para os autores, o DNA não é a única fonte de espaçadores, e a atividade de RT garante memórias de RNA. Mas o que essa descoberta significa? Para as bactérias, a aquisição direta de espaçadores de RNA é uma maneira de ter como alvo parasitas com genoma na forma de RNA. Outra possibilidade é gerar seletivamente espaçadores de regiões intensamente transcritas dos genomas de DNA invasores através da interferência em DNA e RNA. Dessa forma, seriam muito mais eficazes do que espaçadores gerados aleatoriamente a partir do material genético invasor. De fato, as bactérias com esse sistema são bastante resistentes a invasão de elementos genéticos.

3. MIMIVIRE: CRISPR de vírus gigantes?

3.1 Mimivírus

A partir de um “microrganismo” isolado da água de uma torre de resfriamento na Inglaterra, foram descobertos por pesquisadores franceses, em 2003, os “vírus gigantes” (La Scola et al., 2003). Devido à impossibilidade de amplificar sequências de 16S rDNA e ao fato de infectar amebas, este novo vírus foi denominado Mimivírus (de *mimicking*

microbe). Vários estudos posteriores descreveram uma variedade desses vírus gigantes que apresentam genomas circulares de dsDNA maiores que aqueles de muitas bactérias (1,2 Mpb e até 1.200 genes) (Raoult et al., 2004). Muitos estão associados a amebas (principalmente *Acanthamoeba polyphaga*) e algas, apresentam morfologias características e podem ser observados por microscopia óptica (cerca de 0,75 micrômetro). Além do DNA genômico, foi detectada a presença de RNA, tanto mensageiro quanto transportador, no interior do capsídeo, e muitas proteínas envolvidas em transcrição e tradução. É possível que esses RNAs e proteínas tenham participação importante nas primeiras etapas da infecção, o que os tornariam menos dependentes do hospedeiro. A origem destes vírus gigantes ainda está em discussão, mas alguns fatos são impressionantes. Por exemplo, vírus gigantes infectivos foram isolados de amostras de geleiras datadas de 30.000 anos, revelando, portanto, uma longevidade surpreendente (Legendre et al., 2015).

Estudos de genomas de mimivírus descreveram a presença de genes inesperados, como aqueles envolvidos em tradução, e abriram linhas de estudo de sua participação na origem de organismos eucariotos, assim como na determinação dos limites conceituais entre célula e vírus, já que alguns dos genes descritos eram até então observados somente em células. Além desses genes, os estudos mostraram que esses mimivírus podem ser infectados por virófagos, que são pequenos vírus com genoma circular de dsDNA, geralmente da família *Lavidaviridae*, que dependem e parasitam vírus gigantes da família *Mimiviridae* (La Scola et al., 2008). Um dos virófagos mais bem caracterizados é denominado Zamilon (Gaia et al., 2014). Curiosamente foram encontradas linhagens de mimivírus resistentes à infecção pelos virófagos, o que levou à descoberta de um sistema de imunidade genômica denominada MIMIVIRE (*mimivirus virophage resistance element*), que seria análogo ao sistema CRISPR-Cas9 (Levasseur et al., 2016).

3.2 O mecanismo MIMIVIRE

A resistência dos mimivírus aos virófagos seria devido à presença de um sistema de defesa baseado no reconhecimento de sequências específicas de um determinado material genético exógeno (virófago). Esse sistema estaria alicerçado em três sequências contidas no genoma dos mimivírus que compõem os elementos caracterizados como MIMIVIRE: R349 (sequência de repetição que participaria do reconhecimento do DNA-alvo), R350 (possível helicase) e R354 (possível nuclease).

A sequência R349 é composta por uma região de 28 nucleotídeos idênticos à sequência do virófago Zamilon, sendo parte (15 nucleotídeos) repetida quatro vezes (figura 1). A presença dessas quatro repetições idênticas se mostra extremamente relevante para que ocorra o processo de defesa, pois outras linhagens do mesmo vírus que possuem apenas uma repetição não apresentam resposta imune ao virófago Zamilon.

As proteínas codificadas por R350 e R354 catalisariam o desenrolamento e a clivagem do DNA exógeno, respectivamente. Elas atuariam em conjunto com R349 no processo de neutralização do virófago.

O possível sistema de defesa MIMIVIRE apresenta semelhanças com o sistema imune CRISPR-Cas9 das bactérias. Resumidamente, o sistema CRISPR necessita de dois componentes principais para que ocorra a clivagem do DNA exógeno: a proteína Cas9, que é caracterizada por apresentar domínios de nuclease, e o RNA-guia, que, por sua vez, exibe uma sequência específica de 20 nucleotídeos que auxiliam o sistema a localizar o gene-alvo

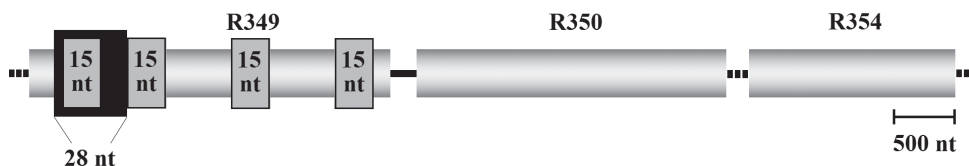


Figura 1. Esquema representativo de MIMIVIRE no genoma dos Mimivírus. Estão esquematizadas as sequências R349, R350 e R354, sistema de defesa contra virófagos. A região de 28 nt, em preto, apresenta identidade com sequências do genoma do virófago Zamilon, e parte dessa é repetida três vezes em R349. As sequências R350 e R354 provavelmente codificam uma helicase e uma nuclease, respectivamente, que complementarizam o sistema de recombinação. Imagem: Oliveira, TC e Schrank A (2016).

pela complementaridade de bases. Dessa forma, o sistema MIMIVIRE poderia ser utilizado como o sistema CRISPR-Cas9 para introduzir alterações genômicas nos virófagos.

Com a futura elucidação do mecanismo de atividade do MIMIVIRE, especialmente na compreensão da função das proteínas que seriam codificadas nas sequências R350 e R354, bem como do mecanismo de integração e endereçamento dessas proteínas pela sequência R349, poderão ser desenvolvidas novas técnicas de edição gênica, assim como ocorre com a tecnologia desenvolvida a partir do sistema endógeno CRISPR.

MIMIVIRE pode vir a fundamentar uma complementação à técnica de CRISPR, pois a endonuclease presente em MIMIVIRE poderia ser aplicada em sequências ou genomas ricos em AT, devido ao fato de a sequência R354 apresentar maior atividade de degradação em moldes de DNA com porcentagem baixa de conteúdo GC (28-38%) (Levasseur *et al.*, 2016). Além disso, o sistema CRISPR necessita de uma sequência PAM, aspecto que não se aplicaria a MIMIVIRE. Portanto, com um sistema desenvolvido a partir de MIMIVIRE, seria possível expandir o número de sítios-alvos.

4. Tópicos especiais sobre CRISPR

Finalizar uma obra sobre uma técnica tão recente e cujas diversas novas aplicações e variações brotam ininterruptamente é um desafio. Creio que, talvez, a maneira mais adequada de encerrar este livro seja mencionar alguns trabalhos curiosos e/ou mais recentes que, em conjunto, sinalizam claramente que estamos diante de um imenso *iceberg* que está emergindo diante de nós: CRISPR.

Podemos citar, por exemplo, a descoberta de proteínas virais que são supressoras do sistema CRISPR (Bondy-Denomy *et al.*, 2015), que podem vir a se tornar uma ferramenta molecular interessante para a regulação final da edição genética. Outro exemplo interessante se refere aos “antimicrobianos programáveis”, *i.e.*, Cas9 e gRNAs entregues a bactérias virulentas resultando na destruição de plasmídeos que conferem resistência a antibióticos (Bikard *et al.*, 2014; Citorik *et al.*, 2014). Outro conceito elegante e muito inovador é o “aparato de gravação molecular” (*molecular recording device*) baseado na aquisição de sequências espaçadoras no sistema endógeno CRISPR, visando o registro de eventos celulares ao longo da história populacional e/ou evolutiva de bactérias (Shipman *et al.*, 2016). Uma estratégia com objetivos semelhantes, porém aplicada a eucariotos, é a GESTALT (*genome editing of synthetic target arrays for lineage tracing*) (McKenna *et al.*, 2016).

Dentre algumas novas aplicações podemos mencionar: (i) protocolos escalonáveis para a marcação de proteínas endógenas com GFP (Leonetti et al., 2016), (ii) mapeamento em larga escala de sequências regulatórias (Rajagopal et al., 2016), (iii) mapeamento em alta resolução de proteínas *in vivo* (Mikuni et al., 2016) e (iv) protocolos para a introdução de mutações homozigóticas ou heterozigóticas específicas (Paquet et al., 2016).

Novos estudos abordando relações inesperadas entre CRISPR e transposons podem auxiliar na elucidação das origens evolutivas deste sistema de defesa bacteriano antiviral (Kapitonov et al., 2015; Krupovic et al., 2016). Por outro lado, a identificação de outras nucleases do sistema CRISPR, como a C2c2, cujo alvo é o RNA em vez de DNA, pode resultar em diversas aplicações originais para a edição, controle e/ou rastreamento de RNAs (Abudayyeh et al., 2016). Por fim, a fusão de Cas9 com outras enzimas pode promover edições ainda mais refinadas, como a estratégia de “*base editing*” (Komor et al., 2016; Nishida et al., 2016), na qual a citidina deaminase ligada à nuclease permite a conversão específica de citidina em uridina (C→T) sem o uso de DNAs moldes e sem envolver clivagem do alvo.

5. Conclusões

Conforme visto ao longo de toda esta obra, a genética molecular está vivenciando uma nova fase: a edição genômica de precisão acessível a todos. A vasta maioria dos dados apresentados nestes 12 capítulos foram gerados em estudos muito recentes, *i.e.*, a partir de 2012. Dia após dia, novas descobertas são feitas, novos produtos são lançados, novas aplicações da técnica são divulgadas. Presenciamos o nascimento de uma nova era da genética, cujos desdobramentos estamos apenas começando a contemplar. O futuro promete nos apresentar um *fantástico mundo novo*.

Referências

- Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*. 2016 Jun;aaf5573.
- Barrangou R. CRISPR-Cas system and RNA-guided interference. *Wires RNA*. 2013;4:267-78.
- Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, et al. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat Biotechnol*. 2014 Nov;32(11):1146-50.
- Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, et al. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature*. 2015 Oct;526(7571):136-9.
- Citorik RJ, Mimee M, Lu TK. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat Biotechnol*. 2014 Nov;32(11):1141-5.
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pírzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 2011;471(7340):602-7.
- Gaia M, Benamar S, Boughalmi M, Pagnier I, Croce O, Colson P, et al. Zamilon, a novel virophage with Mimiviridae host specificity. *PLoS One*. 2014;9(4):e94923.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable Dna-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012;337(6096):816-21.
- Kapitonov VV, Makarova KS, Koonin EV. ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs. *J Bacteriol*. 2015 Dec;198(5):797-807.
- Karoui ME, Baudet V, Schbath S, Gruss A. Characteristics of Chi distribution on different bacterial genomes. *Res Microbiol*. 1999;150:579-87.
- Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 Apr;533(7603):420-4.
- Krupovic M, Shmakov S, Makarova KS, Forterre P, Koonin EV. Recent Mobility of Casposons, Self-Synthesizing Transposons at the Origin of the CRISPR-Cas Immunity. *Genome Biol Evol*. 2016 Jan;8(2):375-86.
- La Scola B, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, et al. A giant virus in amoebae. *Science*. 2003 Mar;299(5615):2033.
- La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, et al. The virophage as a unique parasite of the giant mimivirus. *Nature*. 2008;455:100-4.

- Legendre M, Lartigue A, Bertaux L, Jeudy S, Bartoli J, Lescot M, et al. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. *Proc Natl Acad Sci USA*. 2015 Sep;112(38):E5327-35.
- Leonetti MD, Sekine S, Kamiyama D, Weissman JS, Huang B. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc Natl Acad Sci USA*. 2016 Jun;113(25):E3501-8.
- Levasseur A, Bekliz M, Chabrière E, Pontarotti P, La Scola B, Raoult D. MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature*. 2016 Mar;531(7593):249-52.
- Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*. 2015;520(7548):505-10.
- Marakova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct*. 2006;1:7.
- Marakova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*. 2015;13:722-36.
- Marraffini LA, Sontheimer EJ. RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*. 2010;11(3):181-90.
- Marraffini LA. CRISPR-Cas immunity in prokaryotes. *Nature*. 2015;526:55-61.
- McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole organism lineage tracing by combinatorial and cumulative genome editing. *Science*. 2016 May;aaf7907.
- Mikuni T, Nishiyama J, Sun Y, Kamasawa N, Yasuda R. High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing. *Cell*. 2016 Jun;165(7):1803-17.
- Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, Tabata M, Mochizuki M, Miyabe A, Araki M, Hara KY, Shimatani Z, Kondo A. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*. 2016 Aug 4. "First Release Science Papers" DOI: 10.1126/science.aaf8729.
- Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*. 2016 May;533(7601):125-9.
- Peng W, Feng M, Feng X, Liang YX, She Q. An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res*. 2014;43:406-17.
- Rajagopal N, Srinivasan S, Kooshesh K, Guo Y, Edwards MD, Banerjee B, et al. High-throughput mapping of regulatory DNA. *Nat Biotechnol*. 2016 Feb;34(2):167-74.
- Raoult D, et al. The 1.2-megabase genome sequence of Mimivirus. *Science*. 2004;306:1344-50.
- Samai P, Pyenson N, Jiang W, Goldberg GW, Hatoum-Aslan A, Marraffini LA. Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. *Cell*. 2015;161:1164-74.
- Shipman SL, Nivala J, Macklis JD, Church GM. Molecular recordings by directed CRISPR spacer acquisition. *Science*. 2016 Jun;aaf1175.
- Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, Lambowitz AM, Fire AZ. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science*. 2016;26(6276):aad4234-1-12.
- Sontheimer EJ, Marraffini LA. CRISPR goes retro. *Science*. 2016;351(6276):920-1.

Glossário

Alelo nulo: alelo não funcional.

Amplicon: conjunto de moléculas derivadas da amplificação do DNA-alvo via PCR.

Blunt end: extremidades abruptas. Oposto de extremidades coesivas ou livres (*overhanging*).

Cas9: endonuclease do sistema CRISPR Tipo II responsável pela clivagem do DNA-alvo. Ela é dirigida por um RNA-guia até a sequência-alvo, na qual realiza cortes nas duas cadeias de DNA.

Cas9 HNH⁺: é um tipo específico de *nickase*. Trata-se de uma Cas9 na qual apenas o domínio HNH é cataliticamente ativo (RuvC foi inativado).

Cas9 RuvC⁺: é um tipo específico de *nickase*. Trata-se de uma Cas9 na qual apenas o domínio RuvC é cataliticamente ativo (HNH foi inativado).

Cassete de expressão: construção genética artificial composta pelo gene de interesse e seus elementos regulatórios (promotor e terminador).

cis: refere-se a interações e/ou propriedades na mesma molécula. Interações em *cis* são do tipo intramoleculares.

Clivagem: quebra de ligação covalente na cadeia de DNA ou RNA.

Constitutivo: com expressão ao longo de todo tempo.

Cpf1: nuclease do sistema CRISPR-Cpf1, que é um sistema CRISPR do Tipo V classe 2. A sigla também pode se referir ao próprio sistema CRISPR das bactérias *Prevotella* sp. e *Francisella novicida*.

CRISPR: sistema imune adaptativo antiviral endógeno de bactérias e arqueas. A sigla também pode se referir a um loco bacteriano composto por “repetições palindrômicas pequenas regularmente espaçadas e agrupadas” de função imune. Por fim, ela também pode se referir à técnica de biologia molecular capaz de mediar diversos eventos tais como: edição genética, clivagem ou rastreamento de RNAs, mapeamento genético e regulação da expressão gênica.

D10A: é um tipo específico de Cas9 HNH⁺. Trata-se de uma Cas9 na qual houve uma substituição de aspartato para alanina (D10A) no domínio RuvC I, convertendo a endonuclease em um *nickase*.

Decâmero: composto ou molécula formados por dez unidades de um mesmo monômero.

Degron: sinal de degradação. Elemento associado à cadeia polipeptídica que é suficiente para o reconhecimento por uma maquinaria de proteólise celular.

DNA homólogo: o mesmo que DNA doador.

Dúplex: composto por dois elementos. Dúplex de RNA: composto por duas cadeias de RNA.

Epissomo: material genético extracromossômico que pode se replicar de maneira autônoma ou se integrar no cromossomo.

Frameshift: mudança da fase de leitura. Refere-se a mutações que alteram a forma com que as trincas do mRNA são lidas pelo ribossomo, resultando na alteração da sequência primária do polipeptídeo ou em cadeias truncadas.

Genética Reversa: forma de análise genética cuja direção é “do DNA para o fenótipo”, isto é, gera-se uma mutação no DNA e se observa a alteração fenotípica resultante. Por sua vez, na *Genética Direta* (“do fenótipo para o DNA”) observa-se uma alteração fenotípica na natureza e se procura pela mutação responsável por ela.

H1: um promotor da RNA polimerase III.

H840A: é um tipo específico de Cas9 RuvC⁺. Trata-se de uma Cas9 na qual houve uma substituição de histidina para alanina (H840A) no domínio HNH convertendo a endonuclease em um *nickase*.

Heterodímero: complexo formado por duas partes diferentes.

Heterodúplex: composto por dois elementos diferentes.

Heterozigoto composto: indivíduo portando alelos com mutações distintas.

HNH: um dos domínios catalíticos da Cas9. Ele é responsável pela clivagem da cadeia de DNA que se pareia com a sequência-guia do gRNA.

Homodímero: complexo formado por duas partes iguais.

in planta: na planta viva.

Knock-in: estratégia que permite a integração ou substituição de elementos genéticos (*e.g.*, transgenes, novos alelos) em regiões predeterminadas do genoma.

Knockout: noucaute.

Loco (do latim, locus): localização física de determinado elemento genético no genoma.

MIMIVIRE: sistema de defesa de vírus gigantes contra virófagos (vírus parasitas de vírus) recentemente identificado e que apresenta similaridades com CRISPR.

Molde homólogo: o mesmo que DNA doador.

Motivo (do inglês, motif): pequena sequência nucleotídica (ou de aminoácidos) conservada e que possui função biológica.

Multiplex CRISPR: utilização da técnica de CRISPR para edição de vários locos simultaneamente.

Multiplexing: tornar múltiplo algum evento ou processo.

Mutação in frame: mutação que não resulta na alteração da fase de leitura.

Mutação missense: mutação que resulta na troca de um aminoácido por outro na cadeia polipeptídica.

Mutação monoalélica: mutação que ocorreu em apenas um dos alelos.

Mutação nonsense: mutação que resulta na troca de um aminoácido por um códon de terminação.

Mutações bialélicas: mutações (não necessariamente idênticas) que ocorreram nos dois alelos.

Mutagenic Chain Reaction (reação mutagênica em cadeia): estratégia que permite editar com elevada eficiência os dois alelos de um loco, resultando em mutações homozigóticas.

Nick: clivagem em apenas uma das duas cadeias do dsDNA.

Nickase: enzima que realiza a clivagem em apenas uma das duas cadeias do dsDNA.

Nocaute: inativado. Refere-se ao gene cuja mutação resultou na perda de sua função.

Nocautamento gênico: inativação do gene.

Off-target: fora do alvo.

Oligômero: composto ou molécula formados por algumas (poucas) unidades de um mesmo monômero.

PAM: motivo adjacente ao protoespaçador. Trata-se de uma pequena sequência nucleotídica (e.g., NGG) próxima à *sequência-alvo*, porém na cadeia oposta do DNA, essencial para o início da interação de Cas9 com o DNA-alvo.

PAMmer: oligonucleotídeo apresentador de PAM. Refere-se a uma pequena molécula artificial de DNA contendo o motivo PAM e que é capaz de se parear com o RNA-alvo de fita simples. Esta interação forma uma pequena região de dupla fita portando PAM, característica essa essencial para permitir a interação de Cas9 com o RNA-alvo.

Písceo: relativo a peixe.

RNA-guia: molécula responsável por direcionar a nuclease Cas9 até o DNA-alvo. Ela é composta por duas partes – a sequência-guia e a sequência universal. Trata-se de uma construção quimérica artificial idealizada para simplificar a técnica de CRISPR, a partir da fusão do crRNA e tracrRNA.

RuvC: um dos domínios catalíticos da Cas9. Ele é responsável pela clivagem da cadeia de DNA que se não pareia com a sequência-guia do gRNA.

Scaffold: o mesmo que sequência universal.

Seed: região de aproximadamente 10 nucleotídeos internos da *sequência-guia*, responsável direta pelo início do pareamento do RNA-guia com o DNA-alvo.

Sequência-alvo: região de aproximadamente 20 nucleotídeos no DNA-alvo à qual a *sequência-guia* se pareia.

Sequência-guia: região de aproximadamente 20 nucleotídeos da extremidade 5' do RNA-guia, responsável pelo pareamento com o DNA-alvo.

Sequência universal: região de aproximadamente 80 nucleotídeos da extremidade 3' do RNA-guia, responsável pela ativação da Cas9.

Sinal de localização nuclear (NLS): uma pequena sequência de resíduos de aminoácidos (~10 aa) com papel de direcionar a proteína para o núcleo.

Split-Cas9: duas subunidades não funcionais de Cas9 que, após serem reunidas, podem executar a função original da Cas9 selvagem.

Substituição alélica: troca da sequência nucleotídica do alelo endógeno por outra sequência exógena ligeiramente diferente, portando polimorfismos ou mutações.

Tag: etiqueta. Refere-se a uma molécula ou composto que é adicionado ao DNA ou proteína para marcá-los.

Tetrâmero: composto ou molécula formados por quatro unidades de um mesmo monômero.

Trans: referente a interações e/ou propriedades em outra molécula. Interações em *trans* são do tipo intermoleculares.

Transcrito: molécula gerada pela transcrição gênica, RNA.

Trímero: composto ou molécula formados por três unidades de um mesmo monômero.

Truncado: cortado, mutilado. Refere-se a versões menores e geralmente não funcionais de genes ou proteínas.

TTTTTT: sinal de terminação para a RNA polimerase III. O mesmo que T₆.

U6: um promotor da RNA polimerase III.

Ubíquo: com expressão em todos os tecidos.

Índice remissivo

A

attB: 77

B

Bialélico: 179, 178, 179

BLESS: 13, 142, 144, 145, 146

C

Cas9n: 13, 44, 57, 58, 110, 137

ChIP-seq: 13, 105, 122, 123, 142, 143, 146

Constitutivo: 163-165, 168, 210, 212, 245

Cpf1: 13, 36, 105, 106, 245

CrEdit: 13, 79

CRISPRa: 13, 130, 141, 203, 204, 205, 206

CRISPRi: 13, 61, 63, 101, 130, 141

crRNA: 13, 31, 33, 40, 41, 48, 98, 99, 156,
159, 164, 169, 209, 238

D

D10A: 98, 182, 185, 245

dCas9 ou “dead Cas9”: 13, 57-59, 60-64, 68,
100-105, 110, 114, 122, 123, 130, 137,
143, 146, 148, 171, 172

DECKO: 13, 127, 128, 183

Degron: 245

Digenome: 13, 142, 146

E

EMC: 14, 138-140

ESC: 14, 185, 187

eSpCas9: 14, 103, 148

F

FnCas9: 14, 76

FokI: 14, 89, 90-92, 102, 103, 148

Frameshift: 47, 121-123, 164, 246

G

GUIDE-seq: 14, 142-144, 146

GWAS: 14, 182

H

H840A: 98, 246

Hammerhead: 14, 164, 200

hCas9: 14, 111, 113, 182, 185

Heterozigoto composto: 147, 246

HNH⁺: 13, 98, 100, 245

HTGTS: 14, 142, 144, 146

I

IDAA: 14, 138, 141

IDLV: 14, 144

iPSC: 14, 76, 77, 178, 182

L

LIC: 15, 125

lncRNA: 15, 61, 63, 183

loxP: 88, 89, 211

M

MAMA-PCR: 15, 156

MIMIVIRE: 15, 25, 239-241

Missense: 123, 185, 246

“Mutagenic Chain Reaction” ou MCR: 15,
78, 148, 202, 207-209, 215, 224-229, 246

N

Nickase: 13, 44, 47, 58, 98, 100, 102, 113-115, 182, 185, 245, 246

NmCas9: 15, 101, 121

Nonsense: 123, 246

O

ODM: 15, 223

P

PAMmer: 15, 63, 64, 68, 247

R

RCas9: 16, 68

RFLP: 16, 170, 178

RGEn: 16, 36, 166

RISC: 16, 86, 87

RNAi: 16, 25, 31, 34, 37, 61, 63, 86-89, 93, 123, 183, 195-197, 211, 229

RNP: 16, 147, 170, 184-186

RT:Cas1: 16, 239

RuvC⁺: 13, 98, 100, 245, 246

S

SaCas9: 16, 101, 121

Scaffold: 40, 42, 111, 119, 125, 128, 155, 247

“Sinal de localização nuclear” ou NLS: 15, 111-115, 117, 164, 165, 167, 182-185, 247

SNP: 16, 47, 55, 121, 174, 175

SpCas9: 16, 40, 43, 76, 98, 100-103, 105, 111, 113, 121, 148, 168, 198, 209

SpCas9-HF1: 16, 103, 148

Split-Cas9: 103, 247

St1Cas9: 16, 101, 121

T

T₆: 11, 182-185, 198, 199, 248

TALEN: 16, 86, 89-93, 125, 166, 179, 221

tracrRNA: 17, 40, 41, 48, 98, 99, 106, 156, 159, 160, 164, 169, 209, 247

TSS: 17, 59-62

U

U6: 11, 80, 111, 113, 118, 121, 125, 127, 128, 168, 169, 171, 182-185, 198, 210, 248

Ubíquo: 212, 248

USPTO: 17, 230-232

W

WES: 17, 142, 146

WGS: 17, 142, 146

Z

ZFN: 17, 86, 89-93, 166, 179, 221